Über mögliche Interdependenzen zwischen Fernsehverhalten einerseits und von Lehrerinnen und Lehrern beurteilter Mathematikleistung andererseits bei Schülerinnen und Schülern der Sekundarstufe I

eine sowohl methodische als auch
inhaltlich-didaktische Reflexion

Dissertation
zur
Erlangung des Doktorgrades
Dr. rer. nat.

der Fakultät für
Mathematik
an der

Universität Duisburg-Essen

vorgelegt von
Michael Fey

aus Unna
Januar 2014

Gutachter: Prof. Dr. Gerhard Herden
Gutachterin: Prof. Dr. Katja Lengnink
Gutachter: Prof. Dr. Gerhard Arminger

Vorsitzende des Prüfungsausschusses: Prof. Dr. Lisa Hefendehl-Hebeker

Tag der mündlichen Prüfung: 25. Juni 2014 für Mama und Papa

Mein Dank gebührt in erster Linie meinem Doktorvater Prof. Dr. Gerhard Herden, der mich sowohl in den erfolgreichen Phasen als auch in den kritischen Momenten immer unterstützt hat. Sie haben mir nicht nur eine unfassbare Chance gegeben, sondern mir vor allem durch Ihre warmherzige Betreuung bei der Umsetzung meines Traums geholfen. In vielen anregenden Gesprächen mit Ihnen habe ich nicht nur zahlreiche mathematische und didaktische Kenntnisse hinzugewinnen können, sondern auch viel über die Welt, die Literatur und die Kunst gelernt. Ich danke Ihnen besonders für Ihr Vertrauen und dass Sie mir auch in der Zeit nach der Promotion helfen, in der akademischen Welt meinen Platz entsprechend meiner Fähigkeiten und Qualifikationen zu finden.

Bedanken möchte ich mich aber auch bei Frau Dr. Monika Meise und insbesondere bei Herrn Dr. Uwe Albrecht. Sie haben mich schon zu frühen Zeiten meines Lehramtsstudiums in großem Maße gefördert und dadurch in mir den Wunsch nach einer wissenschaftlichen Laufbahn geweckt.

Eine Dissertation, zumindest besagt dies meine Erfahrung, ist allerdings nicht nur ein universitäres Produkt, geschrieben im dunklen Kämmerlein. Ohne die Unterstützung und Liebe von den Menschen außerhalb des Forschungskreises ist eine solche Arbeit nur schwer zu realisieren. In erster Linie gebührt mein Dank daher meinen Eltern und meinem Bruder, die mich mit vollster Zuneigung und viel Verständnis bei jeder Lebens- und Berufsentscheidung bestärkt und mir jeden Richtungswechsel ermöglicht haben. Dankbar bin ich darüber hinaus auch meiner Freundin Stephanie Ernst und meinen besten Freunden Oliver Zurnieden, Manuela Vida und Sandra Dressler, die jederzeit mit Rat, Trost, Hilfe und neuen Sichtweisen für mich da waren. Danke!

Abschließend bedanke ich mich recht herzlich bei den an der empirischen Untersuchung teilnehmenden Schulen für die immer freundliche und hilfsbereite Mitarbeit.

Abstract

About possible interdependencies between television habits on the one hand and the mathematics performance of students aged 10 to 15 rated by teachers on the other hand – a methodical as well as didactical and content-related reflexion –

Current usage data substantiate that television still presents a very dominant medium in childhood as well as in adolescence, in defiance of an increasing media competition situation provoked by the internet and computer games. Accordingly, an effect of this media-related leisure activity also on different areas of education can be assumed. As television is an acoustical and visual medium, it seems self-evident that this medium has a bigger influence on the social behavior and the language development than on the development of numerical and mathematical capacities. Therefore, it is urgently necessary to let local examinations precede a global analysis of the effects of television on the school performance of students. In this mathematical-didactic dissertation, possible connections between the television habits of 232 12- to 14-year-old adolescents and the corresponding mathematics performances will therefore be examined based on a questionnaire. The work is therewith classified within the global topic "Media and education".

While in the known television-critical studies it is usually warned about the effects of an excessive usage time, television advocates often refer to the positive effects of pedagogically characterized television shows. Accordingly, in this work, the average usage time as well as the kind of television consumption are determined. Additionally, the other leisure behavior is surveyed in order to enable a global interpretation of the results. The mathematics performances are measured through interviews with the tutoring teachers.

On the basis of the surveyed data it can be shown with data analytical techniques that the television watching time can be subdivided into four time intervals (little time, average time, a lot of time, excessive time), that five latent factors form the basis of the television habits of adolescents (entertainment, competition & education, primitive entertainment, humorous entertainment, musical entertainment) and that the leisure behavior of the probands can either be characterized as "active/creative" or as "passive/cognitively unambitious".

The results of the analysis point to the fact that especially the time of watching television and the recreational activities of the adolescents are in a direct negative correlation with the assessment of their mathematical capacities. Those who watch a lot of television and do not occupy themselves active and creatively during their leisure time on average receive negative assessments by their teachers. The preferably watched television shows meanwhile hardly let come to mind a connection to the mathematics performances. Only the consumption of primitive entertainment shows is connected with negative mathematical performances. The empirical study therewith joins seamlessly the mass of critical examinations.

In the focus of the dissertation is furthermore the additional aim of creating a real unity of empirical methodology on the one hand and didactical interpretation on the other hand. Resorting to statistical standard methods and letting the to be interpreted result be dictated by the chosen data analytical method shall be avoided imperatively. As in literature and in the known program packages (Excel, SPSS, R) no really adequate methods which are adapted to the level of ordinally scaled data are available, existing methods were refined and programed. On the basis of own evidence of different mathematical formulas it could therewith be proven, among others, that through the dichotomization of the variables with the help of the conceptual analytical measurement theory according to Wille, the loss of information of factor analytical methods practically disappears. Furthermore, an adequate method for the elimination of appearing missing values is developed.

In the framework of this research project, these new methods can be applied to extensive and real data sets for the first time. In doing so, all methods prove to be of value, so that the most substantial augmentation of knowledge of this dissertation can be seen in this field.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1. Einleitung

„Fernsehen mindert den Schulerfolg" (Welt online 2009), *„Macht zu viel Fernsehen Kinder zu Verbrechern?"* (Bild online 2013), *„Fernsehen macht Kinder dick und dumm"* (Bild online 2011).

Schlagzeilen wie diese, die vor allem nach schwachen Ergebnissen bei Bildungsstudien wie TIMSS oder PISA in den deutschen (Boulevard-) Medien veröffentlicht werden, haben einen beträchtlichen Anteil daran, dass Bildschirmmedien wie das Fernsehen (mittlerweile) einen zweifelhaften Ruf erleiden. Dabei wiederholt sich nüchtern betrachtet lediglich die Geschichte, denn immer wenn ein neues Medium auf den Markt kommt, wird es zunächst durch die Kulturkritik negativ beurteilt. Im 19. Jahrhundert, heute nahezu unvorstellbar, wurden selbst Romane ablehnend behandelt, weil befürchtet wurde, dass die geschriebenen Geschichten der Jugend böse Wünsche in den Kopf setzen. Später galt dann das Telefon lange Zeit als *„Faulmacher"*, das die Menschen von persönlichen Konversationen abhalten würde. Auch das Medium Film litt vom ersten Tag an, laut der New Yorker Gesellschaft zur Vermeidung von Gewalt gegen Kinder, unter der Kritik, ein Medium zu sein, welches junge Menschen zu einem Leben in Sünde und Verwirrung verleiten würde (vgl. Horx ohne Jahr (o.J.), S. 1). Letztlich haben all diese Medien aber vor allem ungewohnte Sicht- und Handlungsweisen ermöglicht, wobei gerade die etablierten Medien das jeweilige neue Medium ablehnten.

Ein kurzer Blick in die Geschichte des Fernsehens zeigt, dass das Medium eigentlich zunächst vor allem die Nachrichtentechnik revolutionierte, wobei die Ursprünge des Fernsehens noch vor dem Ausbruch des *Zweiten Weltkriegs* zu suchen sind. Zu einem Unterhaltungsmedium wurde das Fernsehen erst Jahre später. Die technischen Geburtshelfer Ferdinand Braun und Paul Nipkow stammen sogar aus dem 19. Jahrhundert. Nipkow war maßgeblich an der Entwicklung der mechanischen Bildabtastung, die Braun durch ein elektronisches Verfahren, die als *Braunsche Röhre* bekannte Kathodenstrahl-Oszillographenröhre, ersetzen konnte, beteiligt (vgl. Reisinger 2012). Da die Entwicklung in mehreren Ländern parallel verlief und unterschiedliche Formen konkurrierten, benötigte das Fernsehen im Vergleich zum Radio einen wesentlich längeren Zeitraum, um sich in der Gesellschaft zu etablieren (vgl. Plake 2004, S. 27).

Nachdem jedoch die verschiedensten Anwendungsmöglichkeiten erkannt wurden, drängte das neue Medium mehr und mehr ins Zentrum der öffentlichen Aufmerksamkeit. Der erste regelmäßige Fernsehprogrammbetrieb weltweit wurde am 22. März 1935 in den USA ausgestrahlt. In Deutschland profitierten zunächst vor allem die Nationalsozialisten und erlebten (auch) durch das Fernsehen einen ungeahnten Aufstieg, da das neue Medium die Verbreitung von propagandistischem Material erleichterte und beschleunigte (vgl. Reisinger 2012).

Nach dem Ende des *Zweiten Weltkriegs* brach schließlich das wahre Fernsehzeitalter an, obwohl weiterhin vor allem politische Konflikte das Programm prägten. So war das Medium während des *Kalten Krieges* beispielsweise bei der Verbreitung von Informationen behilflich und stellte sich auch zu dieser Zeit als geeignetes Propagandainstrument dar.

Erst mit den nachlassenden Spannungen zwischen Ost und West in den 80er und 90er Jahren und der Ausweitung der Übertragungsmöglichkeiten übernahm das Fernsehen fortwährend die Rolle des Unterhaltungsmediums. Die nach und nach entstehenden privaten Fernsehsender trieben diese Entwicklung in Deutschland zusätzlich voran (vgl. Plake 2004, S. 28). Demnach ist

> *„die gegenwärtige Entwicklung (...) durch grenzüberschreitende Verbreitung, durch wirtschaftliche Konzentration sowie eine Integration von Programmbetrieb und verschiedenen Telekommunikationsdiensten gekennzeichnet"* (vgl. ebd., S. 28).

Für viele Familien gehört das Medium heute so selbstverständlich zum Tagesablauf wie essen, trinken und schlafen. Aktuelle Nutzungsdaten des Medienpädagogischen Forschungsverbunds Südwest (mpfs) belegen, dass das Fernsehen sowohl im Kindes- als auch im Jugendalter, trotz einer durch das Internet und Computerspiele steigenden medialen Konkurrenzsituation, nach wie vor ein äußerst dominantes Medium darstellt: Über 90 Prozent der Kinder und Jugendlichen verbringen regelmäßig viel Zeit vor dem Fernseher (vgl. mpfs 2013a, S. 10 und mpfs 2013b, S. 12).

Zahlreichen Politikern, man denke an Ursula von der Leyen oder Annette Schavan, und prominenten Experten wie Manfred Spitzer ist dieser Zustand ein Dorn im Auge. Sie setzen das Fernsehen seither unter den Generalverdacht, Kinder und Jugendliche dick, einsam, kriminell und vor allem dumm zu machen. Dementsprechend gehen die Kritiker von einer Auswirkung dieser medialen Freizeitbeschäftigung auch auf verschiedene Bildungsbereiche aus. Da das Fernsehen ein akustisches und visuelles Medium darstellt, ist es naheliegend, dass es auf das Sozialverhalten und die Sprachentwicklung einen höheren Einfluss hat, als auf die

Entwicklung numerischer und mathematischer Fähigkeiten. Insbesondere ist nicht einmal ein positiver Einfluss des Fernsehens auf einige Bildungsbereiche auszuschließen. Daher ist es dringend notwendig, lokale (fachbezogene) Untersuchungen einer globalen Analyse der Wirkung des Fernsehens auf die Schulleistungen von Schülerinnen und Schülern vorausgehen zu lassen. Das in dieser Arbeit vorgenommene Forschungsprojekt bezieht sich dabei, in der Didaktik der Mathematik ansässig, auf den möglichen Zusammenhang zwischen den Rechenfähigkeiten von Jugendlichen und ihrem Fernsehkonsum und ordnet sich damit dem Globalthema „Medien und Bildung" unter.

Im ersten ALLGEMEINEN TEIL wird dazu in die Problematik eingeführt. Ziel dieses Abschnitts ist es, den aktuellen Forschungsstand zu umreißen und offene Fragestellungen zu entdecken. Dazu wird zunächst anhand bekannter Studien des Medienpädagogischen Forschungsverbunds Südwest das Fernsehverhalten von Kindern und Jugendlichen ausführlich beschrieben, um zu verdeutlichen, dass das Fernsehen nach wie vor ein äußerst wichtiges und einflussreiches Medium im Kindes- und Jugendalter darstellt.

Inwieweit diese steigende mediale Freizeitgestaltung als mögliche Ursache für Bildungsrückstände angesehen werden kann, steht im weiteren Verlauf im Fokus. Dazu werden sowohl populäre Meinungen als auch vorliegende empirische Befunde zum Thema „Fernsehen und Bildung" vorgestellt. Einen besonderen Reiz verspricht dabei die Gegenüberstellung zweier äußerst extremer Meinungen von MANFRED SPITZER (2012) und STEVEN JOHNSON (2006). Während SPITZER in seinem Buch *Vorsicht Bildschirm* alleine den bloßen Geräten diabolische Fähigkeiten zuschreibt, ist JOHNSON in *Everything Bad is Good for You* von einer Intelligenz fördernden Wirkung des Fernsehens überzeugt. Beide Autoren verteidigen ihre radikalen Thesen teils populistisch, aber stets vor dem Deckmantel der Wissenschaft und einer augenscheinlich handfesten Argumentation. Exemplarisch ausgewählte empirische Studien belegen allerdings auf einer wissenschaftlichen Basis, dass Fernsehnutzung sowohl negative als auch positive Effekte auf die Bildung von Kindern und Jugendlichen haben kann. Die von Fernsehkritikern häufig zitierten Studien *Mediennutzung und Schulleistung* und *Berliner Längsschnitt Medien* des Kriminologischen Forschungsinstituts Niedersachsens (KFN) stellen beispielsweise fest, dass die Schulleistung zwar von vielen weiteren Einflussfaktoren abhängt, das Fernsehen sich aber bei häufiger Nutzung und dem Konsum vorwiegend problematischer Inhalte negativ auf den Schulerfolg auswirken kann. Im Gegensatz dazu unterstreicht das 2005 in Bangladesch durchgeführte *Sesamstraßenexperiment* die positiven Effekte, die pädagogisch geprägte Fernsehsendungen bei regelmäßigem Konsum auf die Entwicklung von Rechen-, Lese-, und Rechtschreibfähigkeiten haben.

Während es an Studien bezüglich der Auswirkungen von diskussionswürdigem Fernsehkonsum auf die Lese- und Rechtschreibfähigkeiten, den Wortschatz oder andere sprachliche Disziplinen nicht mangelt, ist der ebenfalls äußerst bedeutende Fachbereich der Mathematik bisher in den meisten Untersuchungen sträflich vernachlässigt worden. Dabei setzen gerade gute mathematische Leistungen auch Sprachvermögen, Fantasie, Abstraktionsvermögen und räumliche Vorstellungskraft bei Schülerinnen und Schülern voraus. Die Vermittlung all dieser Kompetenzen und Fähigkeiten sprechen die sogenannten Experten dem Fernsehen explizit ab. Daher erscheint es dringend notwendig, sich konkret mit dem möglichen Zusammenhang zwischen Fernsehverhalten und Mathematikleistung auseinanderzusetzen, so dass die Entscheidung getroffen worden ist, mehrere Fragebogenaktionen an verschiedenen Schulen und Schulformen durchzuführen, um das Fernsehverhalten von Jugendlichen detailliert zu erkunden und mit ihren in der Schule gemessenen Mathematikleistungen in Beziehung setzen zu können.

Aus diesem Grund liegt der Schwerpunkt der Dissertation auf einer selbstständig konzipierten, durchgeführten und ausgewerteten empirischen Untersuchung im METHODISCHEN TEIL. Mit dieser Studie sollen **mögliche Interdependenzen zwischen Fernsehverhalten und in der Schule gemessenen Mathematikleistungen von Jugendlichen** erforscht werden. Dabei ist es nicht das Ziel dieser Arbeit, kausale Beziehungen herauszuarbeiten. Es stehen lediglich mögliche Zusammenhänge im Mittelpunkt des Interesses.

Jugendliche, in dieser Arbeit handelt es sich dabei um Schülerinnen und Schüler der siebten Jahrgangsstufe, eignen sich in besonderem Maße für die Untersuchung, da sie zunehmend autonom über ihr Fernsehverhalten entscheiden, während im Kindesalter die Bindung an die Medien weitestgehend von den Eltern gesteuert wird.

Da sich fernsehkritische Studien vor allem auf die durchschnittliche Dauer des täglichen Fernsehkonsums fokussieren und fernsehlobende Untersuchungen häufig pädagogisch geprägte Programminhalte als Forschungsschwerpunkt wählen, aber nur selten beide Aspekte gleichzeitig analysiert werden, soll in dieser Studie eine Synopse beider Variablen stattfinden. Zusätzlich einbezogen werden soll auch das sonstige Freizeitverhalten der an der Studie teilnehmenden Jugendlichen, um eine globale Interpretation und Einordnung der Ergebnisse zu ermöglichen. Im Zentrum der Untersuchung stehen daher die Ermittlung der dem Fernsehverhalten und dem Freizeitverhalten von Jugendlichen zugrunde liegenden latenten Faktoren sowie die Ermittlung der Abweichung der durchschnittlichen Dauer des täglichen Fernsehkonsums der zu befragenden Jugendlichen von der durchschnittlichen Dauer des täglichen Fernsehkonsums ihrer Altersgruppe.

Um Zusammenhänge zwischen den ermittelten Faktoren und Größen sowie die Interpretation der Ergebnisse und ihre Einordnung in den großen Themenbereich „Fernsehverhalten und Bildung" zu ermöglichen, müssen zudem die Mathematikleistungen der Probanden gemessen werden. Bezüglich der Leistungsmessung existiert in der didaktischen Literatur eine kontroverse Diskussion. Während Vertreter der Reformpädagogik beispielsweise fordern, Schulnoten durch Lernberichte zu ersetzen, suchen Befürworter eines wissens- und stoffdominierten Leistungsbegriffs nach wissenschaftlich legitimierten Methoden zur Leistungsmessung (vgl. VOLLSTÄDT 2005, S. 15). Aufgabe dieser Arbeit soll es allerdings nicht sein, didaktisch motiviert in diese Diskussion einzugreifen. Vielmehr soll der Versuch unternommen werden, die Situation der Leistungsmessung mathematisch zu modellieren, um anschließend zu hinterfragen, ob diese Modelle Lösungen zulassen, welche (möglichen) Expertenmeinungen optimal angepasst sind. Zudem wird ein Vorschlag zur konkreten mathematischen Leistungsmessung in diesem Forschungsprojekt vorgestellt.

All diese Aspekte werden im weiteren Verlauf des METHODISCHEN TEILS ausführlich dargestellt und um die Beantwortung von Fragen und Überlegungen wie *„Welche Personen sollen einen Fragebogen erhalten und ausfüllen? Welche Fragen kommen in den Fragebogen? Was passiert mit den ausgefüllten Fragebögen?"* (KIRCHHOFF et al. 2006, S. 13) ergänzt.

Im Fokus der Untersuchung steht aber vor allem das Ziel, eine wirkliche Einheit von empirischer Methodik einerseits und didaktischer Interpretation andererseits zu schaffen. Es soll also zwingend vermieden werden, auf statistische Standardverfahren zurückzugreifen und sich das zu interpretierende Ergebnis von dem gewählten datenanalytischen Verfahren diktieren zu lassen. Zu diesem Zweck führt zunächst ein mathematischer Teil sowohl in die grundsätzliche als auch in die besondere Problematik empirischer Untersuchungen in der mathematikdidaktischen Forschung ein. Da im Verlauf der vorzustellenden Studie einerseits vermehrt empirische Methoden, vor allem Verfahren der Faktorenanalyse, benutzt werden sollen und andererseits in der Literatur und den bekannten Programmpaketen (Excel, SPSS, R) keine *wirklich* adäquaten Methoden zu Verfügung stehen, die dem Niveau ordinalskalierter Daten angepasst sind, sind Vorüberlegungen betreffend der grundsätzlichen Problematik solcher datenanalytischen Verfahren elementar.

Hinsichtlich des Einsatzes der Faktorenanalyse werden drei Aspekte diskutiert. Zunächst ist das Problem der Kompatibilität der beobachteten Daten mit der entsprechenden Korrelationsmatrix und damit ein grundsätzliches Problem der Interpretierbarkeit von Ergebnissen, die bei Anwendung eines faktorenanalytischen Verfahrens erhalten werden können, zu untersuchen. Im Anschluss konzentriert sich die Arbeit auf das Problem der Adäquatheit von Methoden der Faktorenanalyse in Bezug auf den Messbarkeitsgrad der beobachteten Daten. Tatsächlich lässt sich oft nicht sicher sagen, ob die beobachteten Zufallsvariablen auf einer Intervallskala oder Ordinalskala usw. gemessen werden. Schließlich wird noch das Problem der Adäquatheit des der Faktorenanalyse zugrunde liegenden linearen Modells betrachtet. Das Hauptziel dieses Abschnitts besteht in der Entwicklung eines Vorschlags zur Überwindung dieser Probleme in jeder konkreten Situation, mit der man konfrontiert werden kann. Dieser Vorschlag soll selbstredend im Verlauf der eigenen empirischen Studie Anwendung finden und durch ein an die formale Begriffsanalyse angelehntes Verfahren zusätzliche Bekräftigung finden.

Die formale Begriffsanalyse bietet die Erzeugung und Visualisierung der Begriffshierarchien auf einer mathematisch fundierten Basis. Dabei soll in diesem Fall die formale Begriffsanalyse auf beliebige transitive relationale Strukturen erweitert werden.

Einen weiteren großen mathematischen Abschnitt bildet die Analyse sogenannter *missing values* (*fehlender Werte*), die ein häufig anzutreffendes Problem in der multivariaten Statistik darstellen, da die multivariaten Analyseverfahren, wie die Faktorenanalyse, sich grundsätzlich nur auf vollständiges Datenmaterial anwenden lassen. Leider lässt die Qualität der üblichen Standardverfahren zur Beseitigung der *fehlenden Werte* bei ordinalskalierten Datensätzen, die im Verlauf der Studie bezüglich des Fernsehverhaltens und der Mathematikleistung von Jugendlichen überwiegend Anwendung finden sollen, aufgrund diverser Probleme zu wünschen übrig, so dass im Zentrum dieses Abschnitts die Entwicklung eines eigenen Vorschlags zum Umgang mit diesen *missing values* sowie die teilweise Programmierung dieses Verfahrens in Excel stehen soll.

Sinn der durchgeführten empirischen Untersuchungen ist es daher auch, die neuen bzw. weiterentwickelten Methoden und Verfahren an großen und vor allem realen Datensätzen auf ihre Tauglichkeit zu überprüfen, damit am Ende der Arbeit tatsächlich von einem methodischen Erkenntnisgewinn ausgegangen werden kann!

Allgemeiner Teil

2. Das Fernsehverhalten von Kindern und Jugendlichen

2.1 Faszination Fernsehen

Bewegte Bilder, emotionale Musik, spannende Abenteuer, eigentlich kein Wunder, dass das visuelle Medium Fernsehen zahlreiche Kinder fasziniert und diese sehr viel Zeit vor der *Flimmerkiste* verbringen (vgl. Fangrath 2008, S. 5). Doch die Ursachen für den regelmäßigen Fernsehkonsum von Kindern sind äußerst vielfältig. Besonders die kleinen Kinder sind von den vielen bunten und bewegten Bildern, die das Fernsehen liefert, fasziniert, da sie ein großes Bedürfnis nach Tagträumen haben. Das Fernsehen stillt dieses Bedürfnis durch die ständige Abwechslung, Spannung und Unterhaltung. Die Welt, welche die Kinder sich wünschen, ist nur einen Tastendruck entfernt. Dazu benötigen sie nicht einmal körperliche oder geistige Anstrengungen (vgl. mpfs o.J.). Darüber hinaus begeistert es die Kinder, dass sie mit Hilfe der Fernbedienung und der Programmvielfalt in Sekundenschnelle von einem aufregenden Ereignis zum nächsten gelangen können. Zwar befriedigen auch Bücher und Geschichten das Bedürfnis, dem problembeladenen und fordernden Alltag zu entfliehen, doch gelingt es dem Fernsehen wesentlich leichter, die Kinder in seinen Bann zu ziehen, da es mit Geräuschen, Farben und Musik arbeitet. Zudem fasziniert das Fernsehen die Kinder, weil es ihre Bedürfnisse nach Freiheit und Abenteuern stillt. Gerade in den Großstädten sind diese nur noch schwer erlebbar, wobei das Fernsehen als Ersatzbefriedigung fungiert. Zwar werden Spaß und Spannung nur indirekt und aus zweiter Hand erlebt, doch ermöglicht das Fernsehen den Kindern die Identifikation mit Helden und Figuren, die über Eigenschaften wie Stärke und Macht verfügen (vgl. Muck o.J.).

Dabei schauen Kinder ganz anders fern, als Erwachsene dies tun. Im Kindergarten- und Vorschulalter können die Jüngsten beispielsweise die einzelnen Szenen nur aneinanderreihen, kausale Zusammenhänge oder Rückblenden sind dagegen noch zu abstrakt und damit unverständlich. Sendungen für diese Altersgruppe müssen daher auf simple Charaktere zurückgreifen. Dementsprechend erscheint es angemessen, lediglich einfache Muster wie *Gut* und *Böse* zu verwenden.

Wie Wissenschaftler der Bundeszentrale für gesundheitliche Aufklärung herausfanden, können Kindergartenkinder noch nicht zwischen Fiktion und Realität unterschieden. Folglich ist es für sie nur schwer möglich zu begreifen, dass das, was sie sehen, nicht die Umwelt um sie herum ist. Erst im Vorschulalter erkennen die Kinder langsam, dass es sich nur um Geschichten handelt und sie nicht selbst ein Teil der jeweiligen Szene sind. Frühestens zu diesem Zeitpunkt sollten Kinder regelmäßig fernsehen dürfen (vgl. ebd.).

Zwischen sechs und neun Jahren sind die Kinder nicht mehr ihren Gefühlen ausgeliefert. Das bedeutet, dass sie von den Handlungen auf dem Bildschirm nicht mehr so leicht überwältigt werden. Nach wie vor ist ein moderates Erzähltempo der Geschichten für das Verständnis ausschlaggebend. Wichtig ist außerdem, dass die Kinder stabile Beziehungen und positive Erfahrungen in der realen Welt bereits gesammelt haben. Erst ab zehn Jahren können die Kinder schließlich Realität und Fiktion im Film nahezu vollständig trennen (vgl. Humanistische Aktion 2000). Nun schauen sie fast wie Erwachsene fern. Allerdings gilt es zu beachten, dass Kinder unter 13 Jahren sich von dem Gesehenen noch nicht wie Erwachsene distanzieren können (vgl. Muck o.J.).

2.2 Studien des Medienpädagogischen Forschungsverbunds Südwest

Bevor man sich damit auseinandersetzen kann, ob das Fernsehen einen Einfluss auf die kindliche und jugendliche Entwicklung bzw. den Schulerfolg hat, ist es zunächst erst einmal unumgänglich, zu untersuchen, welchen Stellenwert das Medium im Kindes- und Jugendalter tatsächlich besitzt. Dazu bedarf es einer kontinuierlichen Darstellungen aktueller Nutzungsdaten sowie der Erforschung spezieller Einzelbereiche.

Die wichtigste Datenquelle stammt hierbei vom Medienpädagogischen Forschungsverbund Südwest (mpfs), der gemeinsam von der Landesanstalt für Kommunikation Baden-Württemberg (LFK) und der Landeszentrale für Medien und Kommunikation Rheinland-Pfalz (LMK) getragen wird und gemeinsam mit dem Südwestrundfunk (SWR) seit vielen Jahren Studien zum Medienverhalten der jüngeren Generation durchführt. Seit über einem Jahrzehnt stehen zwei Untersuchungsreihen im Mittelpunkt: die KIM- und die JIM-Studie. Diese für ihre jeweilige Altersgruppen repräsentativen empirischen Erhebungen geben Antworten auf aktuelle Fragen zur Medienausstattung, zur Nutzungsfrequenz und zum Stellenwert der verschiedenen Medien im Alltag von Kindern und Jugendlichen. 2011 sind diese Flagschiffe um eine weitere Untersuchung, die FIM-Studie, ergänzt worden, bei welcher die Mediennutzung in der Familie im Zentrum des Interesses steht und methodisch aufwendig erforscht wird (vgl. mpfs 2012, S. 3).

Im Folgenden sollen nun diese drei Erhebungen allgemein vorgestellt werden. Dabei wird explizit auf die Forschungsergebnisse zum Thema „Fernsehen” eingegangen, um zu verdeutlichen, dass das Fernsehen sowohl im Kindes- und Jugendalter als auch innerhalb der Familie nach wie vor ein äußerst dominantes Medium ist.

2.3 Ergebnisse der KIM-Studie

Die KIM-Studie ist ein Langzeitprojekt, welches seit dem Jahre 1999 im Zweijahresrhythmus wiederholt wird, um die stetig wechselnden Rahmenbedingungen des Medienangebots und die damit verbundenen Veränderungen adäquat abbilden zu können. Dabei werden im Rahmen der Basisuntersuchungen sowohl sechs- bis dreizehnjährige deutschsprachige Kinder bezüglich ihres Medienumgangs in einem persönlichen Interview (CAPI) als auch jeweils ein Haupterzieher (meist die Mutter) zu deren Mediennutzung, dem Medienumgang des Kindes und der Medienerziehung in einem Fragebogen befragt (vgl. mpfs 2013a, S. 6).

Die aktuellste Version der Studie, deren Abkürzung KIM für „Kinder + Medien, Computer + Internet" steht, stammt aus dem Jahre 2012. Die Daten wurden zwischen dem 9. Mai und dem 13. Juli desselben Jahres erhoben. Als Grundgesamtheit standen ca. sechs Millionen Kinder zur Verfügung. Aus dieser Menge wurden die 1.220 Zielpersonen repräsentativ mit Hilfe des Quotenverfahrens ausgewählt und zu den Themen „Freizeitaktivitäten", „Themeninteressen", „Medienausstattung", „Medienbindung", „Medienfunktion", „Fernsehen" „Computer- und Internetnutzung", „Einstellungen zu Computer und Internet", „Computerspiele", „Lernprogramme", „Computer und Schule" sowie „Mediennutzung im familiären Kontext" befragt (vgl. mpfs 2014).

Es ergab sich 2012 folgende Soziodemografie der befragten Kinder:

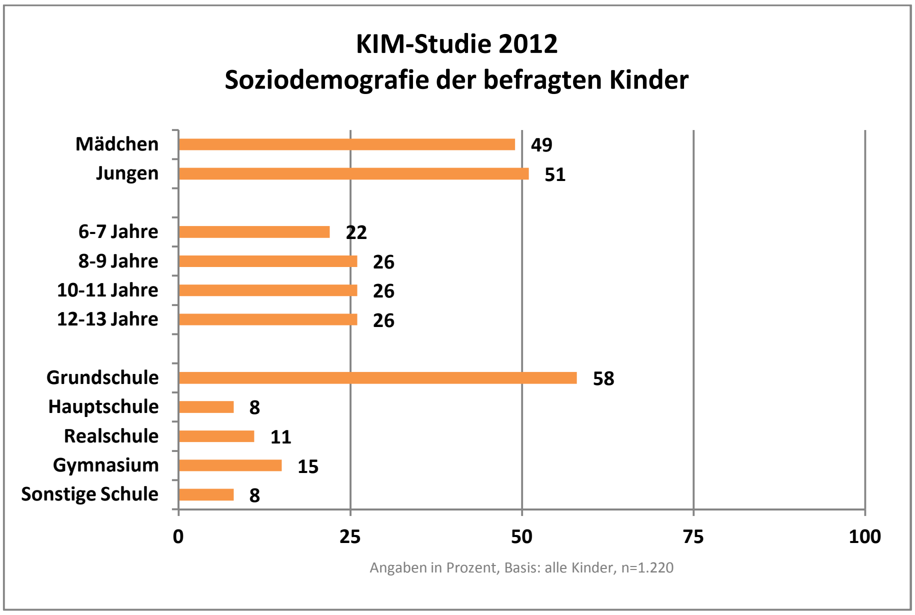


Abbildung 2.1 Soziodemografie der KIM-Studie 2012 (Quelle: KIM-Studie 2012)

2.3.1 Medienausstattung im Haushalt

Zunächst ist bei der KIM-Studie 2012 sowohl die Medienausstattung im Haushalt als auch die Medienausstattung in den Kinderzimmern erforscht worden. Die Ergebnisse, die in Abbildung 2.2 dargestellt werden, zeigen eindeutig, dass in den deutschsprachigen Haushalten, in denen Kinder leben, nahezu eine Vollausstattung mit zahlreichen Mediengeräten vorliegt.

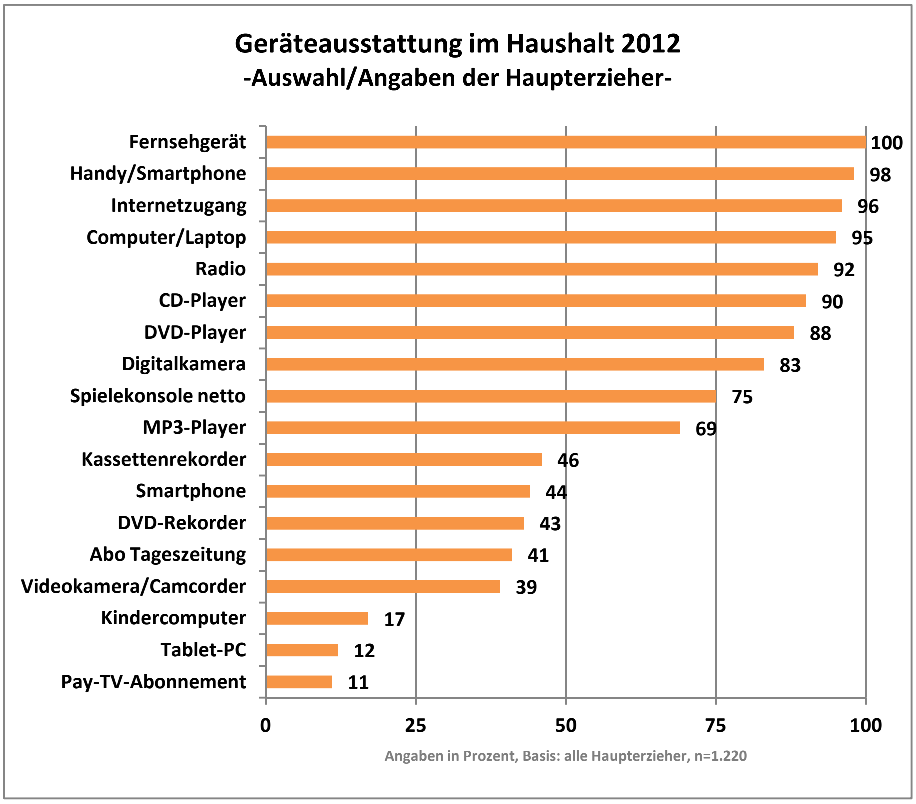


Abbildung 2.2 Geräteausstattung im Haushalt 2012 (Quelle: KIM-Studie 2012)

Dabei handelt es sich vor allem um Fernseher, Handys, Internetzugänge und Computer. Doch auch Radios, CD-Player und DVD-Player sind in etwa neun von zehn Haushalten zu finden. Offensichtlich haben die befragten Kinder alleine durch die Geräteausstattung ihrer Eltern relativ einfach die Möglichkeit verschiedenste Medien zu nutzen.

Noch einfacher wird dieser Zugang natürlich dann, wenn die Kinder in ihren Zimmern sogar eigene Geräte besitzen und nahezu autonom über den Konsum dieser Medien entscheiden können. Laut der KIM-Studie 2012 besitzen bereits 36 % der sechs- bis dreizehnjähren Kinder ein eigenes Fernsehgerät in ihrem Zimmer, wobei die Jungen generell höher ausgestattet sind. Während nur 33 % der Mädchen ein eigenes Gerät besitzen, sind es bei den Jungen nämlich fast 40 %. Der Eigenbesitz nimmt mit steigendem Alter weiter zu, so dass 39 % der zehn- bis elfjährigen Kinder einen eigenen Fernseher beheimaten. Im Vergleich zur Studie von 2010 zeigt sich dennoch ein leichter Rückgang bei der medialen Ausstattung, lediglich der Internetanschluss und der eigene Computer haben demnach in den letzten zwei Jahren zugenommen (vgl. mpfs 2013a, S. 8 ff).

2.3.2 Freizeitaktivitäten und Medienbeschäftigung

Im Rahmen der KIM-Studie ist es 2012 gelungen, die Freizeitaktivitäten der Kinder zu erfragen. Im Mittelpunkt des Interesses steht dabei, welche Rolle die verschiedenen Medien in der schulfreien Zeit spielen.

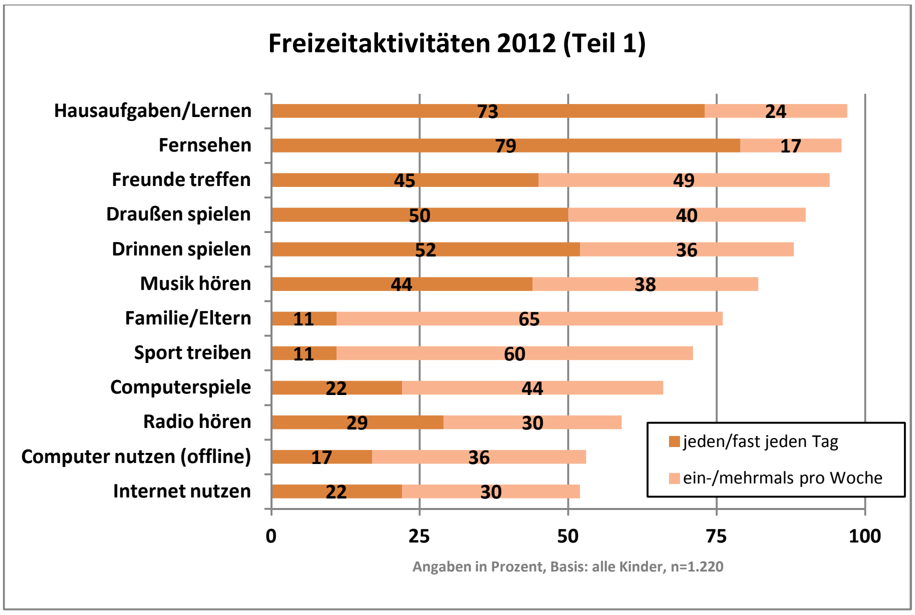


Abbildung 2.3 Freizeitaktivitäten 2012 (Teil 1) (Quelle: KIM-Studie 2012)

Gemäß Abbildung 2.3 nennen fast alle Kinder „Hausaufgaben/Lernen" und „Fernsehen" als diejenigen Aktivitäten, der sie regelmäßig, also mindestens einmal pro Woche, nachgehen. Diese beiden Tätigkeiten scheinen dementsprechend von besonders großer Alltagsrelevanz zu sein. Ähnlich oft treffen sich die Kinder nur mit Freunden bzw. spielen drinnen oder draußen. Andere Medien, man denke an Musik hören, das Spielen von Computerspielen oder die Nutzung des Internets, können in dieser Altersgruppe die Dominanz des Fernsehens (noch) nicht brechen (vgl. ebd., S. 10).

Auch ein Blick auf die tägliche Nutzung zeigt die immer noch immense Bedeutung des Fernsehens bei den befragten Kindern. 79 % geben an, jeden oder fast jeden Tag fernzusehen. Vergleichbare Werte gibt es nur beim Erledigen der Hausaufgaben (79 %). Das Handy wird in dieser Altersgruppe beispielsweise nur von knapp einem Drittel der Kinder täglich genutzt. Die beschriebenen Beobachtungen bleiben altersunabhängig für das Fernsehen, die Hausaufgaben und das Treffen von Freunden stabil. Dahingegen nimmt das Spielen, egal ob drinnen oder draußen, im Laufe der Zeit ab, während Spielekonsolen oder der Computer an Bedeutung gewinnen (vgl. ebd., S. 11 f).

Unabhängig von der Nutzungshäufigkeit vermag die KIM-Studie 2012 auch zu zeigen, wie beliebt das Fernsehen bei Kindern ist. Zwar treffen sich 52 % der befragten Kinder am allerliebsten mit ihren Freunden, doch liegt das Fernsehen hinsichtlich der Frage nach der beliebtesten Freizeitbeschäftigung unangefochten auf dem zweiten Rang (38 %). Mit einem großen Abstand von dreizehn Prozentpunkten folgen Computer- und Videospiele auf Platz drei. Dabei ist das Fernsehen seit 2010 in der Gunst der Kinder sogar um sechs Prozentpunkte gestiegen, so dass tatsächlich davon ausgegangen werden kann, dass die Bedeutung des Fernsehens nach wie vor als besonders groß einzustufen ist (vgl. ebd., S. 12).

2.3.3 Mediennutzung: allein oder in der Familie?

Zwar schauen 47 % der befragten Kinder gerne und regelmäßig gemeinsam mit ihren Eltern oder Geschwistern fern, doch zeigt die KIM-Studie auch, dass immerhin 44 % der Kinder vorzugsweise alleine fernsehen. Dementsprechend integrieren diese Kinder das Medium selbstbestimmt in den eigenen Alltag, bleiben allerdings zu einem Großteil mit ihren Medienerfahrungen alleine und haben niemanden, mit dem sie diese medialen Erfahrungen direkt besprechen können. Mit zunehmendem Alter nimmt die Autonomie der Mediennutzung wenig überraschend zu (vgl. ebd., S. 13 f). Weitere Einzelheiten können Abbildung 2.4 entnommen werden:

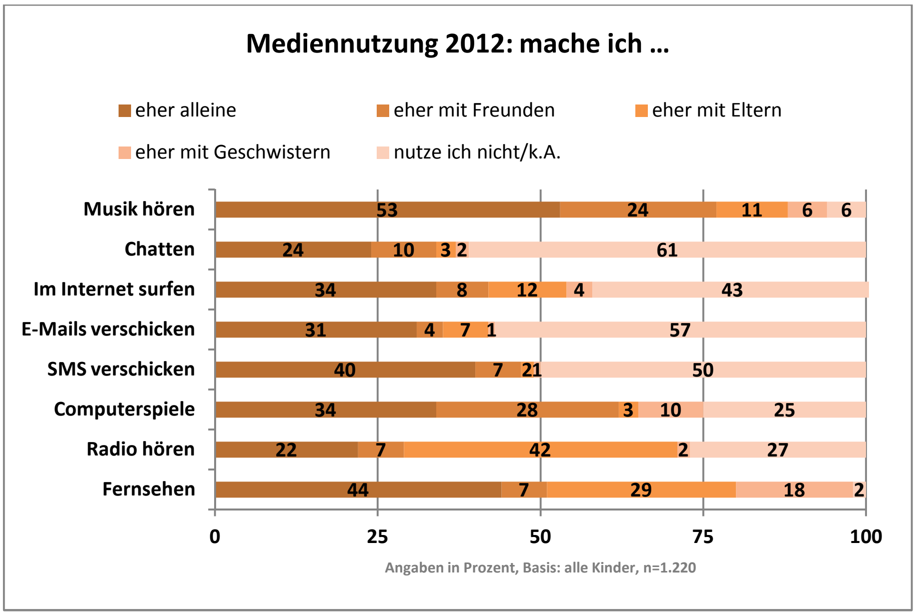


Abbildung 2.4 Mediennutzung 2012 (Quelle: KIM-Studie 2012)

2.3.4 Medienbindung

Jenseits von Nutzung und Besitz wird zudem regelmäßig die subjektive Wichtigkeit der verschiedenen Medien untersucht. Anhand der nachfolgenden Abbildung ist ersichtlich, dass das Fernsehen gerade in jungen Jahren als besonders wichtig eingestuft wird. Altersabhängig geben zwischen 75 % (6-7 Jahre) und 54 % (10-11 Jahre) der Kinder an, nicht auf den Fernseher verzichten zu können. Erst bei den zwölf- bis dreizehnjährigen fällt dieser Wert unter

50 %. Im Zuge dessen gewinnen nun Computer und Internet deutlich an subjektiver Wichtigkeit.

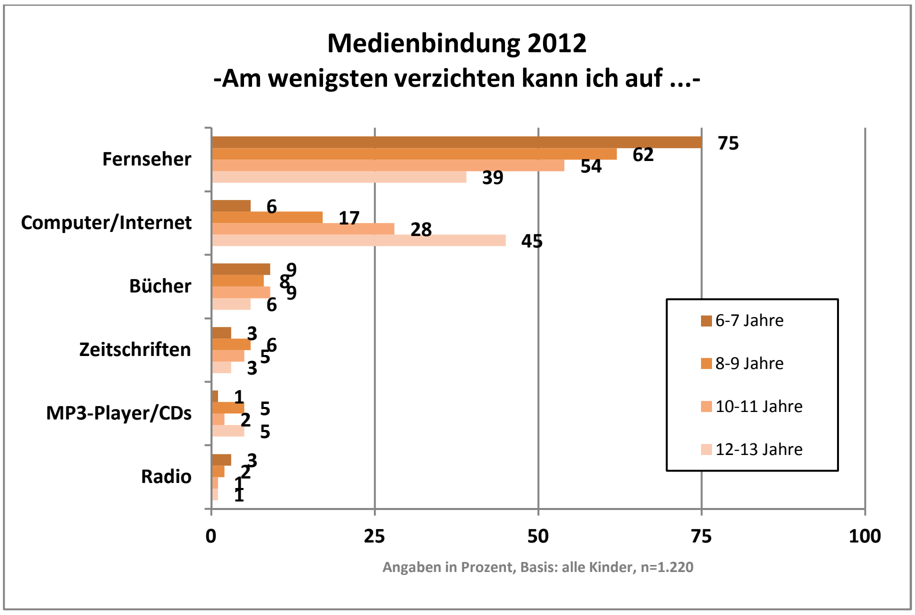


Abbildung 2.5 Medienbindung 2012 (Quelle: KIM-Studie 2012)

Interessanterweise zeigt sich bei dieser Frage ein offensichtlicher Zusammenhang zwischen der Bindung der Haupterzieher an ein Medium und der Bindung des Kindes an das entsprechende Medium. Dies meint, dass, wenn der Haupterzieher nicht auf das Fernsehen verzichten kann, meist auch das Kind eine so enge Bindung zum Fernsehen hat, dass es ebenfalls nicht darauf verzichten mag (vgl. ebd., S. 15 f). Die Bedeutung der Eltern in der Medienerziehung wird bereits an dieser Stelle ersichtlich.

2.3.5 Lieblingsprogramm

Die Macher der KIM-Studie beschäftigen sich des Weiteren mit den inhaltlichen Vorlieben der Kinder. Dabei wurde festgestellt, dass Kinder durchschnittlich mit drei Jahren erste ernsthafte Erfahrungen mit dem Fernsehen machen und fortan relativ schnell ihren eigenen Geschmack entwickeln. Demnach haben insgesamt 82 % der sechs- bis dreizehnjährigen Kinder eine Lieblingssendung. Diese sind sehr vielfältig, so dass die Nennungen eine breite Mischung aus Kinder- und Erwachsenensendungen dokumentieren. Mit zunehmendem Alter haben die Kinder zwar seltener spezielle Lieblingssendungen, diese stammen dann aber, falls vorhanden, aus dem Erwachsenenprogramm (vgl. ebd., S. 18).

Ähnliches gilt auch für die von den Kindern favorisierten Fernsehsender: 60 % haben einen Lieblingssender. Im Kindesalter dominieren dabei vor allem der öffentlich-rechtliche Kinderkanal (26 %) und der Privatsender SuperRTL (26 %). RTL (13 %) und Pro7 (8 %) spielen in dieser Altersgruppe noch keine allzu bedeutende Rolle (vgl. ebd., S. 19).

Diese Beobachtungen werden durch eine weitere Erkenntnis untermauert: Die Kinder sind danach befragt worden, ob sie einen zielgerichteten Grund haben, wenn sie den Fernseher einschalten. Dabei zeigt sich, wie unten stehende Abbildung bekräftigt, dass knapp die Hälfte der Kinder den Fernseher genau dann nutzt, wenn eine bestimmte Sendung, beispielsweise die Lieblingssendung, ausgestrahlt wird. Mädchen verhalten sich etwas häufiger derart zielgerichtet als die Jungen. Mit zunehmendem Alter lassen sich die Kinder allerdings immer häufiger vom Angebot überraschen (vgl. ebd., S. 20).

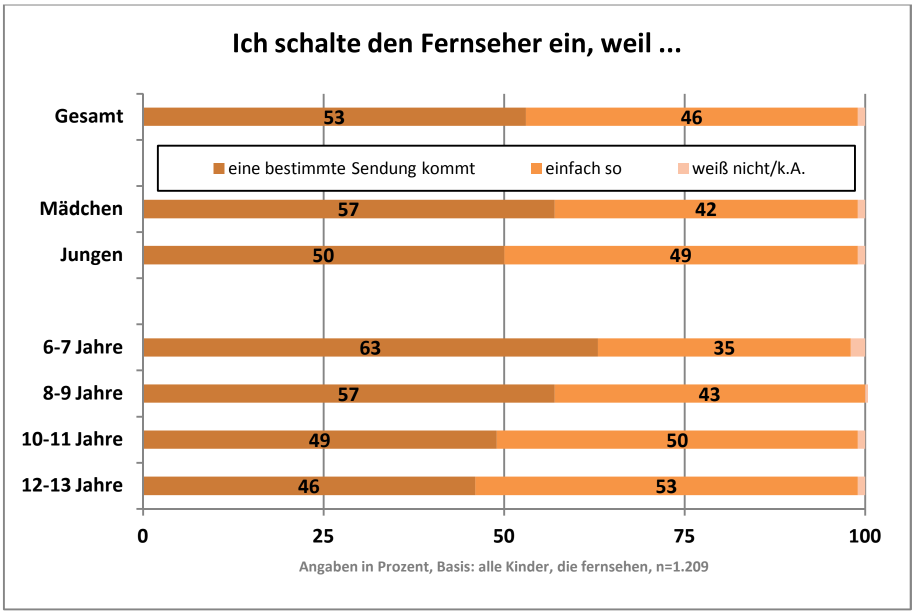


Abbildung 2.6 Ich schalte den Fernseher ein, weil... (Quelle: KIM-Studie 2012)

2.3.6 Problematische Inhalte und Nachrichtensendungen

Um die programmliche Vielfalt des kindlichen Fernsehkonsums auch hinsichtlich problematischer Inhalte noch differenzierter zu betrachten, sind die Kinder in einem ersten Schritt danach gefragt worden, ob sie bereits Erfahrungen mit beängstigenden Fernsehsendungen gesammelt haben. Bei der Betrachtung der Ergebnisse wird deutlich, dass Kinder durchaus selbst erfahren, dass sie von Fernsehinhalten verängstigt werden können. Beispielsweise geben 23 % an, dass sie schon Inhalte gesehen haben, die nicht für sie geeignet waren und 11 % führen an, dass sie beim Fernsehen häufiger mit unangenehmen Bildern konfrontiert werden. Werden die Kinder in einem zweiten Schritt aufgefordert, problematische Inhalte konkret zu nennen, werden vor allem Gruseliges, Gewalt und sexuelle Inhalte beschrieben (vgl. ebd., S. 20).

Interessanterweise haben die Kinder aber teilweise sogar Angst vor den schrecklichen Bildern, die in Nachrichtensendungen gezeigt werden. Insgesamt schauen immerhin 27 % der Kinder regelmäßig Nachrichtenformate. In frühen Jahren handelt es sich vor allem um die Kindernachrichtensendung *logo!* (43 %). Spätestens ab dem zehnten Lebensjahr gewinnen dann aber auch Erwachsenensendungen an Bedeutung (vgl. ebd., S. 21).

2.3.7 Nutzungsdauer

Zu Beginn dieser Ausführungen ist bereits beschrieben worden, dass das Fernsehen von vielen Kindern nahezu täglich genutzt wird. Um diese Feststellung näher zu untersuchen, wurden die Haupterzieher gebeten, die tägliche Mediennutzungszeit ihrer Kinder zu schätzen. Das Ergebnis wird in Abbildung 2.7 dargestellt.

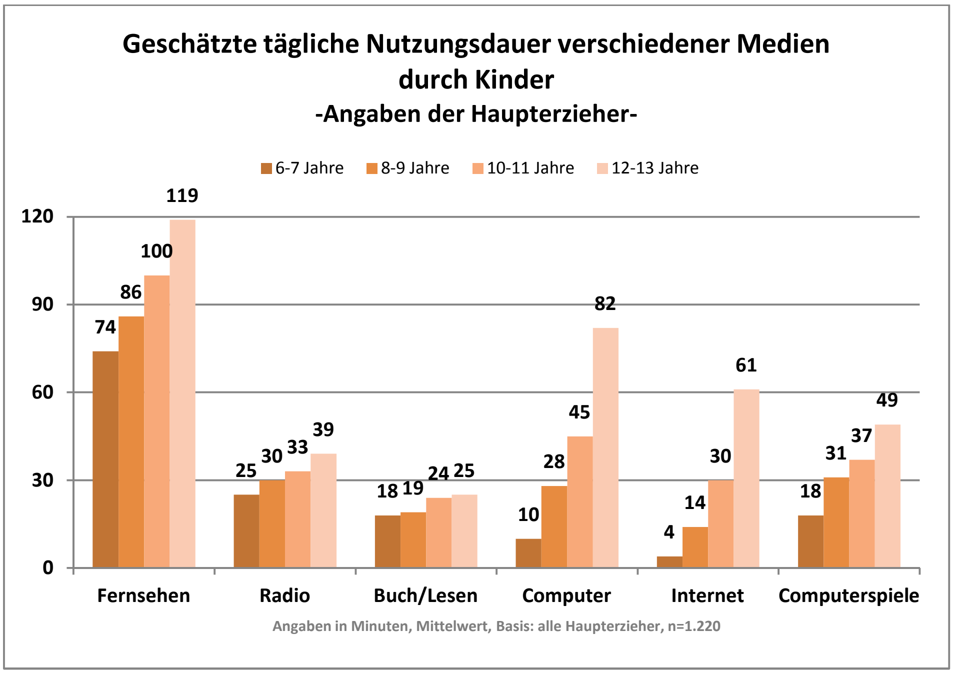


Abbildung 2.7 Geschätzte tägliche Nutzungsdauer verschiedener Medien durch Kinder (Quelle: KIM-Studie 2012)

Auch diese Statistik unterstreicht die Dominanz des Mediums Fernsehen in der Kindheit. Durchschnittlich schauen die Kinder täglich 95 Minuten fern. Der Computer folgt mit einer mittleren Nutzungsdauer von 42 Minuten auf dem zweiten Platz. Auffällig ist, dass alle abgefragten Medien mit zunehmendem Alter länger genutzt werden. Fällt der Anstieg beim Lesen und beim Radio noch recht moderat aus, so nimmt die tägliche Nutzungszeit des Computers und des Internets drastisch zu. Auch beim Fernsehen muss die Zunahme des täglichen Konsums nicht mit der Lupe gesucht werden (vgl. ebd., S. 26 f).

2.3.8 Zusammenfassung

Zusammenfassend deuten die Ergebnisse der KIM-Studie darauf hin, dass das Fernsehen nach wie vor das wichtigste und dominanteste Medium für Kinder im Alter von sechs bis dreizehn Jahren ist, auch wenn die Rollen von Computer und Internet in den letzten Jahren und mit zunehmendem Alter an Bedeutung gewinnen. Das Fernsehen ist die beliebteste Freizeitbeschäftigung und wird fast täglich genutzt. Betrachtet man die gesehenen Inhalte, wird deutlich, dass zwar überwiegend kindgerechte Sender und Sendungen geschaut werden, dies allerdings zum einen altersabhängig ist und zum anderen viele Kinder bereits Erfahrungen mit beängstigenden und für sie verbotenen Inhalten gemacht haben.

2.4 Ergebnisse der JIM-Studie

Ebenfalls ein Langzeitprojekt ist die JIM-Studie, die, seit 1998 jährlich durchgeführt, 2012 ihr fünfzehnjähriges Jubiläum feiern konnte (vgl. mpfs 2013b, S. 3). Im Gegensatz zur KIM-Studie erfasst sie das Medienverhalten der Altersgruppe der zwölf- bis neunzehnjährigen Jugendlichen und steht für „Jugend, Information, (Multi-) Media". Im Rahmen der Basisuntersuchung wurden zwischen dem 7. Mai und dem 17. Juni repräsentativ 1.201 ausgewählte Jugendliche telefonisch befragt (CATI). Die Stichprobe ist aus einer Grundgesamtheit von sieben Millionen deutschsprachigen Jugendlichen ausgewählt worden (vgl. ebd., S. 4).

Dabei ergab sich 2012 folgende Soziodemografie der befragten Jugendlichen:

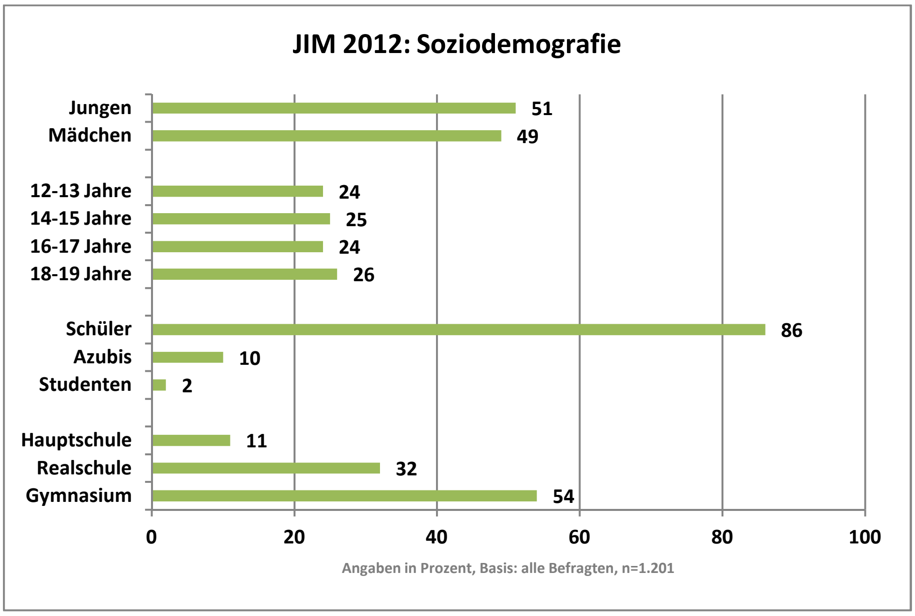


Abbildung 2.8 Soziodemografie der JIM-Studie 2012 (Quelle: JIM-Studie 2012)

2.4.1 Medienausstattung

In Anlehnung an die KIM-Studie erfasst auch die JIM-Studie zunächst die Geräteausstattung im Haushalt im Allgemeinen und in den Zimmern der Jugendlichen im Speziellen. Abbildung 2.9 bezieht sich dabei auf die Geräteausstattung im Haushalt der Jugendlichen.

Es zeigt sich, dass auch in den Haushalten, in denen die zwölf- bis neunzehnjährigen Jugendlichen leben, eine Vollausstattung mit bestimmten Mediengeräten besteht, die den befragten

Jugendlichen den medialen Zugang erleichtern. Dabei handelt es sich vor allem um den Computer, das Handy, den Fernseher oder den Internetzugang. Auch eine Digitalkamera, ein Radio oder ein MP3-Player ist in den allermeisten Haushalten zu finden (vgl. ebd., S. 6).

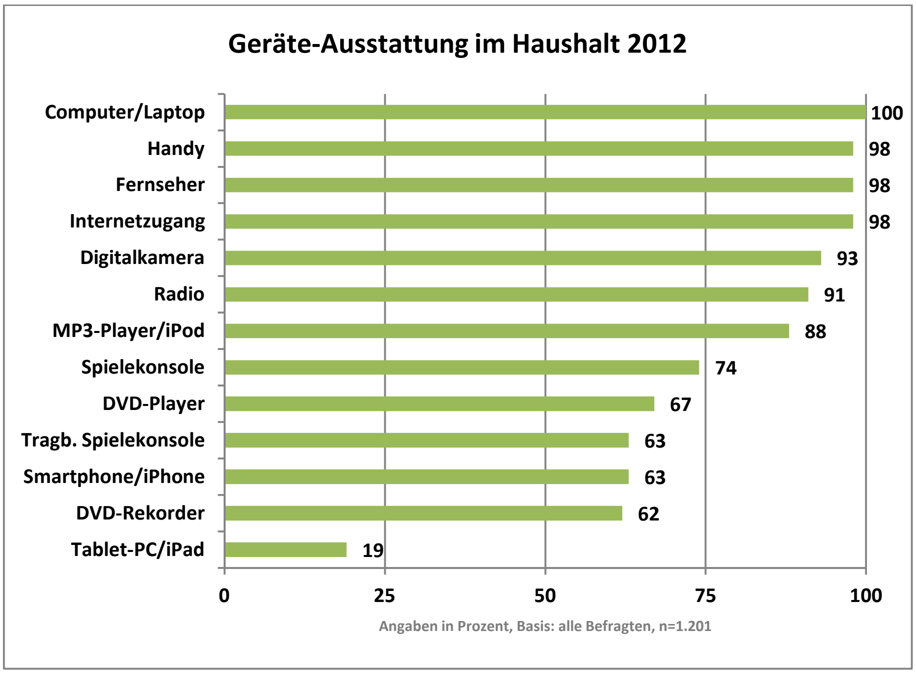


Abbildung 2.9 Geräte-Ausstattung im Haushalt 2012 (Quelle: JIM-Studie 2012)

Unterschiede zu den Ergebnissen der KIM-Studie werden deutlich, wenn man den Gerätebesitz der Jugendlichen aus der JIM-Studie betrachtet. Nahezu alle Jugendlichen verfügen nun über ein eigenes Handy (97 %) und oft auch über einen eigenen Internetzugang (86 %) sowie einen eigenen Computer (82 %). Über einen eigenen Fernseher verfügen allerdings lediglich 55 % der Mädchen und 64 % der Jungen, wobei auch dieser Wert deutlich über dem Vergleichswert aus der KIM-Studie liegt (vgl. ebd., S. 7 f).

2.4.2 Medienbeschäftigung in der Freizeit

Eine zentrale Erkenntnis der KIM-Studie ist, dass das Fernsehen das dominante Medium im Kindesalter ist. Davon kann auf Basis der JIM-Studie im Jugendalter zwar nicht mehr gesprochen werden, doch lässt sich eine mitunter große Bedeutung des Fernsehens auch bei der älteren Altersgruppe nicht leugnen:

So werden laut Abbildung 2.10 Internet, Fernsehen und Handy von jeweils über 90 % der Jugendlichen regelmäßig genutzt. Zusätzlich sind auch Musikmedien wie MP3-Dateien (81 %), das Radio (78 %) oder CDs (56 %) ebenfalls fest im jugendlichen Alltag verankert. Im Vergleich zum Vorjahr sind die erhobenen Daten stabil, so dass sich keine eklatanten Verschiebungen erkennen lassen.

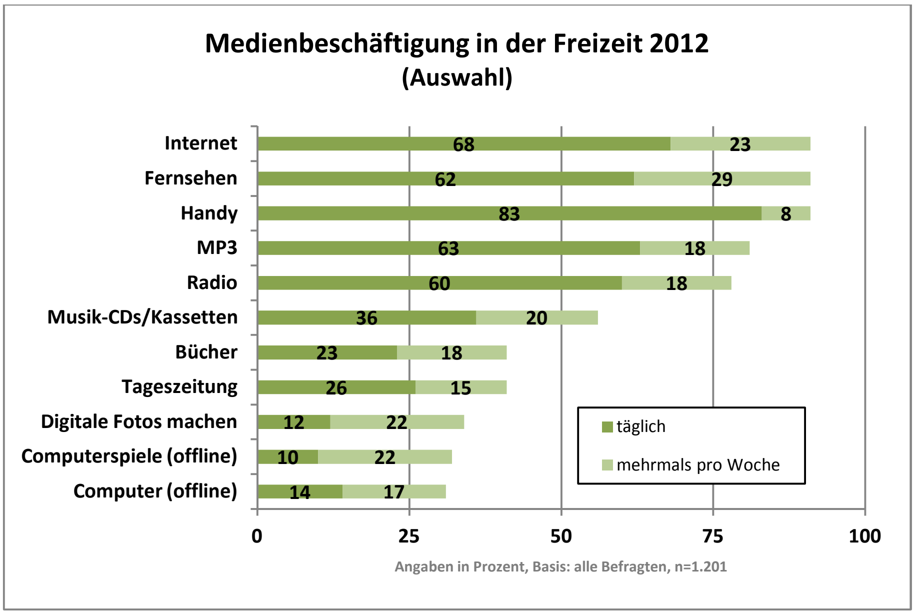


Abbildung 2.10 Medienbeschäftigung in der Freizeit (Quelle: JIM-Studie 2012)

Geschlechtsspezifische Unterschiede beim Fernsehkonsum gibt es in der von der JIM-Studie erfassten Altersgruppe nicht. 90 % der Mädchen und 92 % der Jungen schauen regelmäßig fern. Ebenso bleibt die Regelmäßigkeit der Fernsehnutzung mit zunehmendem Alter konstant (vgl. ebd., S. 12 f).

2.4.3 Wichtigkeit der Medien

Um unabhängig von Gerätebesitz und Nutzungshäufigkeit die Wichtigkeit der Medien für die Jugendlichen zu bestimmen, sind die Probanden gebeten worden, auf einer Skala von „gar nicht wichtig" bis „sehr wichtig" anzugeben, wie viel Bedeutung das jeweilige Medium für sie besitzt.

Bei der Betrachtung des Antwortverhaltens in Abbildung 2.11 zeigt sich, dass das Hören von Musik die wichtigste mediale Tätigkeit im Alltag der Jugendlichen ist. Auf diese Aktivität kann die Altersgruppe scheinbar gar nicht verzichten. Knapp dahinter folgen das Internet und das Handy, welches jeweils 88 % bzw. 80 % der Jugendlichen als „wichtig" oder „sehr wichtig" einstufen. Das Fernsehen wird dagegen lediglich von knapp über 55 % als wichtiges Medium genannt. Die Jungen liegen mit 57 % deutlich vor den Mädchen, welche das Fernsehen nur zu 52 % für (sehr) wichtig halten. Dieser Wert ist damit deutlich geringer als der in der KIM-Studie erfasste Prozentsatz. Mit steigendem Alter ist die Bedeutung des Fernsehens zudem noch weiter rückläufig. Nur noch 48 % der Neunzehnjährigen geben an, nicht auf das Fernsehen verzichten zu können (vgl. ebd., S. 14 f).

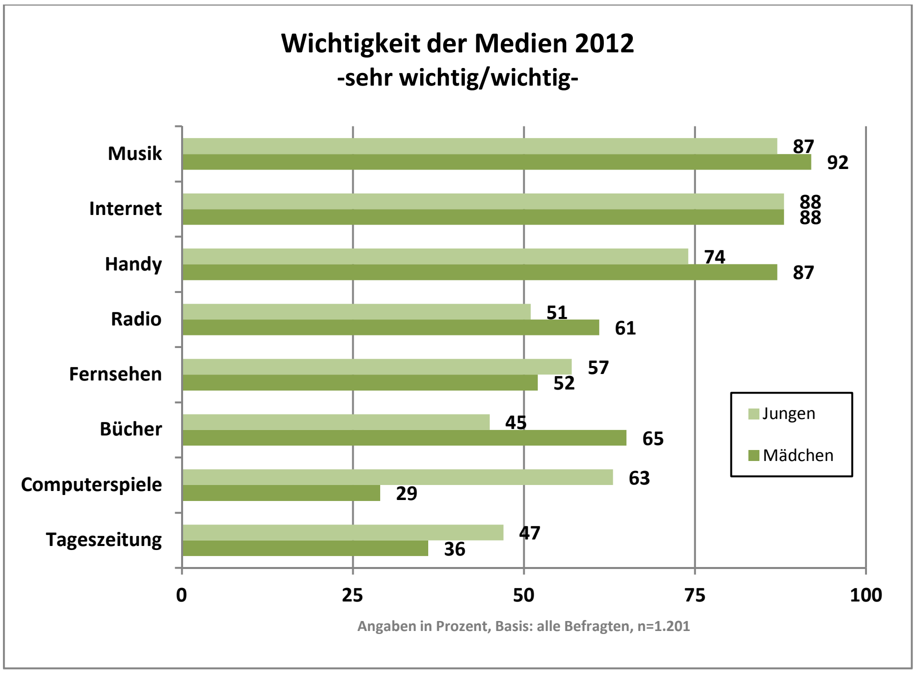


Abbildung 2.11 Wichtigkeit der Medien 2012 (Quelle: JIM-Studie 2012)

2.4.4 Glaubwürdigkeit der Medien

Im Unterschied zur KIM-Studie ist im Verlauf der JIM-Studie auch erfasst worden, inwieweit die Jugendlichen die genutzten Medien für glaubwürdig halten. Dazu wurde untersucht, ob die Jugendlichen die breite Informationsfülle kritisch hinterfragen und welchem Medium sie bei sich widersprechender Berichterstattung am ehesten vertrauen würden. Die passende Frage lautete:

> *„Stell Dir mal vor, Du wirst im Radio, im Fernsehen, in Tageszeitungen oder im Internet über dasselbe Ereignis informiert, die Berichte widersprechen sich aber bzw. sind voneinander verschieden. Wem würdest Du am ehesten glauben: dem Radio, dem Fernsehen, dem Internet oder der Tageszeitung?"* (ebd., S. 16)

Dem Antwortverhalten zufolge trauen die meisten, nämlich knapp die Hälfte der Jugendlichen, unabhängig vom Geschlecht, den Tageszeitungen die seriöseste Berichterstattung zu. Das Fernsehen liegt mit einem deutlichen Abstand von 26 Prozentpunkten immerhin auf dem zweiten Rang. Interessanterweise vertrauen die Jugendlichen damit genau dem Medium am meisten, welches sie eher selten nutzen, während ausgerechnet das beliebte Internet sehr kritisch beurteilt wird. Nur 11 % würden den Onlineangaben glauben. Dieses Ergebnis deutet durchaus darauf hin, dass die Jugendlichen die medialen Angebote kritisch hinterfragen.

Abbildung 2.12 illustriert das Ergebnis:

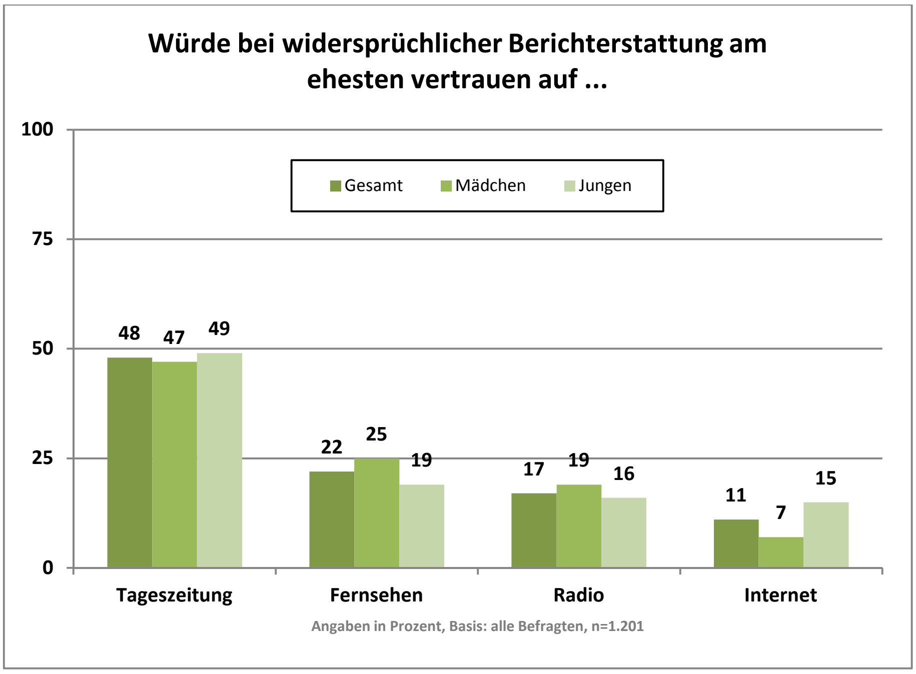


Abbildung 2.12 Glaubwürdigkeit der Medien 2012 (Quelle: JIM-Studie 2012)

Jungen und Mädchen unterscheiden sich vor allem bei der Einschätzung des Fernsehens (eher Mädchen) und des Internets (eher Jungen). Zudem liegt das Fernsehen in der Gruppe der Jugendlichen mit niedrigem Bildungsniveau vier Prozentpunkte vor der Tageszeitung auf Platz eins. Die Tageszeitung verdankt ihre guten Glaubwürdigkeitswerte vor allem den Gymnasiasten (55 %), welche dem Internet kaum trauen (8 %) (vgl. ebd., S. 16 f).

2.4.5 Fernsehnutzung und Lieblingssender

Nach ihrer täglichen Nutzungsdauer gefragt, geben die Jugendlichen an, dass sie durchschnittlich an jedem Wochentag 111 Minuten fernsehen. Dieser Wert ist nur zwei Minuten geringer als im Vorjahr ausgefallen. Deutliche Unterschiede in der Nutzungsdauer lassen sich in Abhängigkeit vom Bildungsniveau feststellen. Demnach schauen Schülerinnen und Schüler der Hauptschule nach eigener Angabe täglich 137 Minuten fern, wohingegen Gymnasiasten das Medium lediglich 96 Minuten nutzen (vgl. ebd., S. 25).

Während laut der KIM-Studie im Kindesalter vor allem die Fernsehsender KIKA und SuperRTL konsumiert werden, steht im Jugendalter Pro7 mit großem Abstand an erster Stelle. 59 % der Jungen und 42 % der Mädchen geben an, dass Pro7 das beste Fernsehprogramm ausstrahlt. Auf dem zweiten Platz befindet sich, vor allem dank der Mädchen, das Programm von RTL, welches von 15 % der befragten Jugendlichen besonders gerne geschaut wird. Die übrigen Sender erhalten kaum Stimmen. Dies gilt auch für die ARD und das ZDF (vgl. ebd., S. 25 f).

2.4.6 Scripted Reality-Formate

In jedem Jahr beschäftigt sich die JIM-Studie explizit mit einem aktuellen Phänomen der Fernsehkultur. 2011 handelte es sich dabei beispielsweise um „Castingshows" wie *Deutschland sucht den Superstar* oder *Germany's next Topmodel*. Da diese Sendungen seitdem in der Zuschauergunst tendenziell abfallen, ist für die aktuelle Untersuchung ein neues Format ausgewählt worden: Scripted Reality-Sendungen. Die Macher der JIM-Studie begründen ihre Entscheidung folgendermaßen:

> *„Der Markt der sogenannten Doku-Soaps ist in den letzten beiden Jahren um eine Facette 'reicher' geworden. Scripted Reality Formate zeigen erfundene Geschichten, die von Laiendarstellern nach einem vorgegebenen Drehbuch (Script) relativ frei umgesetzt werden. Durch spezielle gestalterische Mittel, wie zum Beispiel eine dokumentarisch anmutende Kameraführung, soll beim Zuschauer ein möglichst authentischer Eindruck erweckt werden. […] Es wird befürchtet, dass Kinder und Jugendliche die Machart dieser Sendungen nicht durchschauen (können) und die dokumentarische Aufbereitung dazu führt, dass Fiktion als Realität (miss-) verstanden wird"* (ebd., S. 26 f).

Zur Überprüfung dieser Vermutung wurden zunächst die drei quotenstärksten Formate auf ihre Bekanntheit hin abgefragt. Es handelt sich dabei um die Sendungen *Berlin - Tag & Nacht*, *Verdachtsfälle* und *Familien im Brennpunkt*. Diese Sendungen werden jeweils von fast 40 % der Jugendlichen, vorwiegend von Mädchen, regelmäßig geschaut. Während der Konsum von *Verdachtsfälle* und *Familien im Brennpunkt* mit zunehmendem Alter abnimmt, bleibt *Berlin - Tag & Nacht* auf einem konstanten Level (vgl. ebd., S. 27).

Im zweiten Schritt sind die Jugendlichen, welche die Sendungen regelmäßig sehen, gefragt worden, ob ihnen die Scripted Reality-Formate gefallen. Demnach finden 80 % der Nutzer Gefallen an *Berlin - Tag & Nacht*, während die anderen beiden Sendungen nur von 60 % als gut oder sehr gut eingestuft werden (vgl. ebd., S. 27).

In einem dritten Schritt wurde den Jugendlichen abschließend die fundamentale Frage gestellt, ob sie das Gesehene für real halten. Es zeigt sich, dass tatsächlich nur der Hälfte der Jugendlichen bewusst ist, dass sowohl die Geschichten als auch die Menschen in den Sendungen *Familien im Brennpunkt* und *Verdachtsfälle* frei erfunden sind. 44 % (*Familien im Brennpunkt*) bzw. 45 % (*Verdachtsfälle*) der Jugendlichen sind zwar immerhin davon überzeugt, Schauspieler präsentiert zu bekommen, unterliegen aber dennoch der falschen Annahme, dass es sich um reale Begebenheiten handelt, die von den Schauspielern nachgespielt werden. Ganze 6 % (*Familien im Brennpunkt*) bzw. 4 % (*Verdachtsfälle*) sind sogar dem Irrglauben verfallen, dass in den Doku-Soaps real existierende Menschen in ihrem normalen Alltag gezeigt werden (vgl. ebd., S. 28).

Einen noch interessanteren Einblick bieten die Antworten im Fall von *Berlin - Tag & Nacht*, da die Themen der Sendung noch dichter an die Lebenswelt junger Menschen anknüpfen. So zeigt sich, dass einerseits viermal so viele Jugendliche das Gezeigte für real (16 %) halten als bei den anderen beiden Doku-Soaps. Andererseits sind aber auch wesentlich mehr Urteilende davon überzeugt, dass die Sendung komplett gestellt ist (62 %). Dies lässt sich möglicherweise dadurch erklären, dass das Prinzip der Sendung, es wird das Leben von Menschen zwischen 18 und 42 Jahren gezeigt, die in wechselnden Konstellationen an verschiedenen Orten in Berlin wohnen, einerseits sehr real wirkt, andererseits aber Geschichten wie in *richtigen* Fernsehserien erzählt werden. Betrachtet man abschließend die soziodemografischen Merkmale, fällt auf, dass vor allem Mädchen und Jugendliche mit Hauptschulhintergrund auf die realitätsnahe Darstellung hereinfallen (vgl. ebd., S. 29).

Die Befürchtung, dass die Jugendlichen die Machart der Sendungen nicht durchschauen, findet somit zumindest bei einem nicht zu unterschätzenden Anteil der Probanden Bestätigung.

2.4.7 Zusammenfassung

Wenngleich das Fernsehen als weniger glaubwürdig als die Tageszeitung und als weniger wichtig als das Internet eingestuft wird, hat das visuelle Medium seinen hohen Stellenwert auch 2012 behauptet. Dies lässt sich leicht anhand der bereits beschriebenen Nutzungshäufigkeit und der persönlichen Geräteausstattung begründen. Die JIM-Studie hat zwar einerseits gezeigt, dass die Jugendlichen durchaus in der Lage sind, im Fernsehen und anderen Medien verbreitete Nachrichten kritisch zu hinterfragen, doch wurde gleichzeitig deutlich, dass fast 50 % der untersuchten Altersgruppe nicht dazu fähig sind, Scripted Reality-Sendungen als *Fake* zu identifizieren. Daher muss davon ausgegangen werden, dass die Medienerziehung durch Schule und Familie auch im pubertierenden Alter eine bedeutende Rolle spielt.

2.5 Ergebnisse der FIM-Studie

Nachdem sowohl die Auswertungsergebnisse der KIM- als auch der JIM-Studie in den letzten Jahren eine immer größer werdende Bedeutung medialer Freizeitbeschäftigungen im Kindes- und Jugendalter belegen konnten, wurden diese beiden Basiserhebungen zuletzt mit der FIM-Studie („Familie, Information, Medien") um einen weiteren Blickwinkel ergänzt, indem die ganze Familie in den Mittelpunkt der Untersuchung gerückt wurde. Sinnvoll erscheint diese Differenzierung alleine deswegen, weil die familiäre Medienerziehung und der innerfamiliäre Medienumgang einen spürbaren Einfluss auf das kindliche und jugendliche Medienverhalten vermuten lassen. Der Schwerpunkt der FIM-Studie liegt dabei konkret auf der innerfamiliären Kommunikation, dem Vorhandensein von kommunikativen Verhaltensmustern innerhalb der Familie, Gemeinsamkeiten und Unterschiede der Generationen bei der Mediennutzung sowie der Bedeutung von Medientechnik und Medieninhalten im familiären Gefüge (vgl. mpfs 2012, S. 3).

Um derartig komplexe Einsichten in das Familienleben in Deutschland zu gewinnen, wurde ein besonders aufwendiger Weg beschritten. Zunächst sind aus der Grundgesamtheit von 18 Millionen deutschsprachigen Haushalten mit einem oder mehreren Kindern zwischen drei- und neunzehn Jahren 260 Familien, ausgewählt nach Anzahl und Alter der Kinder, dem Wohnort und der Familienform, in zwei Etappen befragt worden. Dabei fanden, analog zu der KIM-Studie, Face-to-Face-Interviews (CAPI) mit allen Familienmitgliedern statt, mit dem Ziel, Basisdaten für quantitative Statistiken zu erhalten. Zusätzlich wurden Tagebucherhebungen zum Tagesablauf angefertigt (vgl. ebd., S. 4).

Auch im Rahmen der FIM-Studie spielt das Fernsehverhalten eine dominante Rolle. Im Folgenden soll aus diesem Grund explizit auf die entsprechenden Forschungsergebnisse eingegangen werden.

2.5.1 Fernsehen und innerfamiliäre Kommunikation

Im Rahmen der Analyse innerfamiliärer Kommunikationsformen stellt die FIM-Studie fest, dass während des Fernsehkonsums ein Kommunikationsmangel innerhalb der Familie herrscht. Demnach sprechen etwa 50 % der Eltern beim gemeinsamen Fernsehen selten oder nie mit ihren Kindern. Am ehesten geben die zwölf- bis neunzehnjährigen Jugendlichen an, sich beim Fernsehen mit ihren Eltern zu unterhalten (vgl. ebd., S. 25 f).

Wenn in der Familie allerdings über Themen gesprochen wird, *„die mit Mediennutzung, Medientechnik oder Medieninhalten assoziiert sind"* (ebd., S. 38), stellt das Fernsehen den Themenschwerpunkt dar. 58 % der Eltern geben an, dass „Fernsehen oder Dinge, die man im Fernsehen gesehen hat" in der Familie regelmäßig als Gesprächsgrundlage dienen. Damit liegt das Fernsehen in dieser Kategorie beispielsweise vor der Tageszeitung. Die FIM-Studie begründet diese Erkenntnis damit, dass das Fernsehen und die Fernsehinhalte aus Sicht vieler Eltern eher als Mode, Musik und Sportereignisse zum Alltag der Befragten gehören, so dass sich kontinuierlich neuer Gesprächsstoff ganz von alleine entwickelt (vgl. ebd., S. 38).

Die Angaben der Eltern werden durch das kindliche Antwortverhalten gestützt, schließlich geben auch 66 % der Kinder und Jugendlichen an, dass sie häufig mit ihren Eltern über das Fernsehen betreffende Themen reden oder diskutieren, wobei die Häufigkeit mit zunehmendem Alter ein wenig nachlässt (vgl. ebd., S. 40).

Es konnte belegt werden, dass sowohl die Eltern untereinander über Themen des Fernsehens kommunizieren als auch der generationsübergreifende Austausch stattfindet. Die elterninternen Gespräche gehen dabei nahezu gleichermaßen häufig von den Müttern und den Vätern aus. Allerdings findet die Eltern/Kind-Kommunikation über das im Fernsehen Gesehene eher zwischen den Kindern und den Müttern statt (vgl. ebd., S. 48).

Detaillierte Informationen bzgl. der Kommunikationsrichtungen beim Thema „Fernsehen bzw. was man darin gesehen hat" liefert dazu die nachfolgende Abbildung:

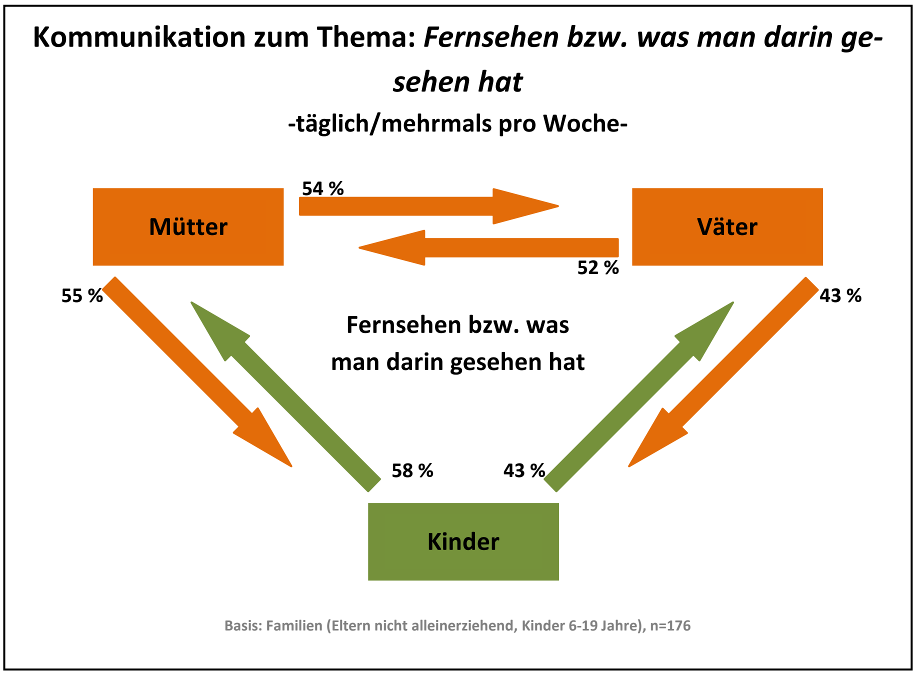


Abbildung 2.13 Kommunikation zum Thema „Fernsehen bzw. was man darin gesehen hat" (Quelle: FIM-Studie 2011)

2.5.2 Fernsehdauer in der Familie

Dem Fernsehen wird aber nicht nur in der innerfamiliären Kommunikation mehr Aufmerksamkeit geschenkt als den anderen Medien, auch die tägliche Nutzung unterstreicht die dominante Position des Fernsehens im gemeinsamen Familienalltag. Dies meint, dass 27 % der Eltern täglich gemeinsam mit ihren Kindern fernsehen. Eine ähnliche Relevanz kommt lediglich dem Radio (21 %) zu (vgl. ebd., S. 63). Bei den Angaben der Kinder und Jugendlichen wird die herausstechende Stellung des Fernsehens als Familienmittelpunkt noch offensichtlicher, da 80 % der Grundschüler und 69 % der Jugendlichen regelmäßig mit ihren Eltern fernsehen (vgl. ebd., S. 65).

Basierend auf den Tagebucheinträgen der Eltern wurden die gemeinsamen Medientätigkeiten für das Fernsehen, das Radio, das Internet und das Lesen von Büchern im Tagesverlauf ermittelt. Während Bücher und Internet scheinbar zu keinem Zeitpunkt am Tag, wenn überhaupt gegen 19:00 Uhr vor dem Zubettgehen der Kinder, eine bedeutende Rolle im gemeinsamen Familienalltag spielen, wird das Radio den ganzen Tag über relativ konstant von 10 % der Eltern als gemeinsame mediale Aktivität hervorgehoben. Bei der Fernsehnutzung ist über den ganzen Tag verteilt kaum eine gemeinsame Nutzung vorhanden. Erst am Abend nimmt das gemeinsame Fernsehen rasant zu. Den offensichtlichen Hochpunkt erfährt die Kurve gegen 20:00 Uhr. Knapp 20 % der Eltern nutzen zu dieser Zeit das Fernsehen zusammen mit ihren Kindern (vgl. ebd., S. 67 f).

Die quantitative Basiserhebung untermauert die Tagebucheinträge. Demnach geben sowohl die Eltern als auch die Kinder an, vor allem beim Abendessen und vor dem Zubettgehen gemeinsame Zeit vor dem Fernseher zu verbringen (vgl. ebd., S. 67 f).

2.5.3 Inhaltliche Fernsehnutzung in der Familie

Den Themenkomplex „Fernsehen" abschließend, beschäftigt sich die FIM-Studie mit der inhaltlichen Fernsehnutzung in der Familie. Dabei geben die Eltern an, dass sie vor allem Kindersendungen (11 %), Zeichentrick (9 %), Seifenopern/Telenovelas (8 %), Quiz- und Spielshows (7 %) sowie Castingshows und Krimis (je 5 %) mit ihren Kindern schauen. Sind die Kinder noch jung, werden vor allem Kindersendungen (28 %) und Zeichentrick (26 %) gesehen. Ab dem Grundschulalter lässt die Bedeutung solcher Sendungen für den gemeinsamen Fernsehkonsum nach (nur noch 13 % und 7 %). Wichtiger werden nun Seifenopern/Telenovelas (11 %), Quiz- und Spielshows (6 %) und Wissenschaftssendungen (7 %). Im Jugendalter stehen dann eher Krimis (10 %) auf dem gemeinsamen Fernsehprogramm (vgl. ebd., S. 75). Die gemeinsame Sendungsauswahl scheint dementsprechend durchaus altersadäquat zu sein.

2.5.4 Zusammenfassung

Die FIM-Studie belegt, dass das Fernsehen nicht nur für Kinder und Jugendliche von besonderer Wichtigkeit ist. Auch im Kreis der Familie ist es das dominante Medium, was vor allem anhand von zwei Aspekten belegt werden kann. Zum einen wird über kein anderes mediales Thema derart häufig innerhalb der Familie kommuniziert. Dies zeigt, wie tief sich das Fernsehen im Alltag der Familie verankert hat. Zum Anderen wird kein anderes Medium vergleichsweise häufig gemeinsam genutzt. Gerade am Abend schafft es das Fernsehen noch immer, die Familie zu versammeln. Zudem verdeutlichen die Auswertungsergebnisse erfreulicherweise, dass zumindest bei der familiären Mediennutzung die inhaltliche Programmauswahl altersadäquat abläuft.

In ihrer Gesamtheit unterstreichen die drei vorgestellten Erhebungen des Medienpädagogischen Forschungsverbunds Südwest die Popularität des Fernsehens. Trotz attraktiver Alternativen (z.B. Smartphones, Internet) kann es seinen Platz in den Kinder-, Jugend- und Wohnzimmern als zumindest wichtiges Freizeitmedium behaupten.

3. Fernsehen und Bildung

3.1 Öffentliche Meinungen aus Politik, Fernsehen und Kultur

Praktisch zeitgleich mit der Ausstrahlung des ersten Fernsehprogrammbetriebs verschafften sich auch die ersten kritischen Stimmen Gehör. Bereits 1961 beschrieb Newton Minows, der damalige Vorsitzende des Federal Communications Commitee, in seiner Rede „Television and the Public Interest" das Fernsehprogramm als *„Aneinanderreihung von (...) Blut und Donner, Aufruhr, Gewalt, Sadismus und Mord"* (zitiert nach Johnson 2006, S.74). Besondere Bedeutung gewinnt diese kritische Aussage alleine dadurch, dass die TV-Landschaft zur damaligen Zeit eigentlich eher von anständigen Entertainern wie Andy Griffith und Perry Como und weniger von großen Skandalsendungen dominiert wurde. An der Grundeinstellung vieler Politiker und anderer öffentlicher Personen hat sich in den letzten 50 Jahren allerdings wenig geändert. Ganz im Gegenteil, die Vorurteile und Vorwürfe haben sogar zugenommen (vgl. Johnson 2006, S. 74). Aus diesem Grund soll im folgenden Abschnitt ein grober Überblick über derartige Meinungen aus der Öffentlichkeit gegeben werden.

Besonders deutlich fällt das Urteil von Ursula von der Leyen (2010) aus. Im ZDF-Polit-Talk „Berlin Mitte" stellt die damalige Familienmisterin der CDU fest: *„Wir wissen alle, dass das Fernsehen dick, dumm, traurig und gewalttätig macht"* (zitiert nach Prantl 2010). Mit dieser eindeutigen Aussage, die leider argumentativ nahezu unbegründet bleibt, erstickt sie jede mögliche andere Meinung und aufkommende Diskussion im Keime. Leider unterlässt sie es damit aber auch, zwischen verschiedenen Programminhalten und Zuschauern zu differenzieren. So wirkt ihr Urteil unreflektiert und einseitig.

Auch Marcel Reich-Ranicki (2008a), der einst sogar die „Goldene Kamera", einen anerkannten Fernsehpreis der Zeitschrift HÖRZU, ablehnte, beschreibt den Einfluss des Mediums Fernsehen auf die Bildung drastisch, indem er sagt: *„Fernsehen macht die Dummen dümmer und die Klugen klüger"* (zitiert nach Stern online 2008). Die Aussage des bis zu seinem Tod einflussreichsten deutschen Literaturkritikers der Gegenwart beinhaltet einerseits, dass die *„Dummen"* sich vorzugsweise auf besonders anspruchsarme Fernsehkost fokussieren, während andererseits die *„Klugen"* so intelligent wie er seien und nahezu keine Zeit vor dem Fernseher verbringen oder sich zumindest mit anspruchsvollen Sendungen die Zeit vertreiben. Das wird deutlich, wenn er hinzufügt:

> *„Man kann im ARTE-Programm manchmal sehr schöne, wichtige Sachen sehen. Ich habe früher auch wichtiges im 3-Sat-Programm gesehen, aber das hat sich heute geändert, meist kommen da schwache Sachen, aber nicht der Blödsinn, den wir heute* [Anm.: Reality TV u.ä.] *sehen"* (REICH-RANICKI 2008b).

Er betont also, dass das Fernsehen die intellektuellen Fähigkeiten des Einzelnen verstärkt und damit die Schere zwischen intelligenten und weniger intelligenten Menschen immer weiter auseinander treibt. ELKE HEIDENREICH (2008), die ehemalige Moderatorin der ZDF-Literatursendung *Lesen!* bekräftigt die von REICH-RANICKI (2008a und 2008b) formulierte Fernsehkritik auf drastische Art und Weise. Mit den Worten *„Wie jämmerlich unser Fernsehen ist, wie arm, wie verblödet, wie kulturlos, wie lächerlich"* (zitiert nach TZ-online 2008) unterstreicht sie, dass das Fernsehen in seiner heutigen Form ihrer Meinung nach keine Bildung vermittelt.

Das TV-Urgestein THOMAS GOTTSCHALK (2012) greift bei einem Vortrag an der Heidelberger Universität zu der Fragestellung, ob sich Quote und Anspruch in der heutigen Fernsehunterhaltung noch auf einen Nenner bringen lassen, die angeführte Problematik auf. Für GOTTSCHALK ist es dabei heutzutage fast unmöglich, mit einem qualitativ und intellektuell hochwertigen TV-Programm ein Millionenpublikum zu erreichen (zitiert nach CROLLY 2012). Daher ist es nicht verwunderlich, dass der von REICH-RANICKI (2008b) lobend hervorgehobene TV-Sender ARTE sich nur geringer Einschaltquoten erfreuen kann, während das vielerorts als „Hartz IV"-Fernsehen diffamierte Programm von RTL und Co. von deutlich mehr Menschen geschaut wird.

Die TV-Kritikerin KLAUDIA WICK schlägt bei ihrem Vortrag während der Münchener Medientage 2012 eine ähnliche Richtung ein und gibt den Zuschauern sogar eine gewisse Mitschuld an dem Zustand des (deutschen) Fernsehens:

> *„Das Publikum hat eine folgenreiche Entscheidung getroffen. Es möchte vor dem Bildschirm nicht mehr gefesselt werden. Es möchte lieber alte Schinken noch einmal sehen, als sich die neuen Serienfiguren draufzuschaffen. Es möchte lieber intellektuell unterfordert, als emotional durchgeschüttelt werden. Es sucht mit dem Finger auf der Fernbedienung nicht nach Herausforderung, sondern nach Beiläufigkeit"* (WICK 2012).

Nach WICK (2012) bildet das Fernsehen die Menschen also nicht, doch sind die Individuen daran sogar selbst schuld. Den Grund dafür sieht sie in einer Abwertung der *„Aufmerksamkeitsökonomie"* des Fernsehens. Die Fernsehkritikerin sieht im Fernsehen heutzutage ein

„Restzeitmedium", welches erst dann eingeschaltet wird, *„wenn wir für alles andere zu müde, zu zerstreut, zu ausgepowert sind"* (ebd.).

Auch der polnische Satiriker Wieslaw Brudzinski (o.J.) ist von den negativen Folgen des Fernsehens überzeugt. Er kritisiert, dass den Kindern und Jugendlichen durch den überhöhten Fernsehkonsum die Kreativität und Fantasie verloren gehen würden. Dies wird deutlich, wenn er Fernsehen als die *„... Phantasie der Armen"* (Brudzinski o.J.) beschreibt und damit die geistig Armen meint. Er weicht damit deutlich von der Vorstellung einiger Fernsehbefürworter ab, die anmerken, dass *Lücken und mehrdeutige Szenen* in modernen Fernsehsendungen den Zuschauer durchaus dazu auffordern, sich bestimmte Dinge vorzustellen und mit eigenen Bildern und Ideen zu füllen.

Der deutsche Autor und Satiriker Dieter Hildebrandt (o.J.) glaubt ebenfalls, dass das Fernsehen, zumindest bei exzessiver Nutzung, zu gesundheitlichen Risiken und Bildungsrückständen führt. *„Notorische Fernsehzuschauer bekommen größere Gefäße und immer kleinere Köpfe"* (zitiert nach Livenet o.J.), schlussfolgert er.

Comedian Kaya Yanar (o.J.) fügt in Anlehnung an den hohen Fernsehkonsum der Kinder und Jugendlichen scherzhaft hinzu, dass deutsche Schülerinnen und Schüler bei Tests in der Schule sogar den Telefonjoker, bekannt aus Quizsendungen wie *Wer wird Millionär*, verlangen würden und unterstreicht damit indirekt den großen Einfluss, den das Fernsehen mittlerweile auch auf andere, wichtigere Bereiche des jugendlichen und kindlichen Lebens gewonnen hat (zitiert nach Quotez.net o.J.).

Andere prominente Persönlichkeiten, die das Fernsehen stark kritisieren, fürchten, dass durch das visuelle Medium zusätzlich sprachliche Fähigkeiten und die familiäre Gemeinschaft beeinträchtigt werden. Werner Schneyder (o.J.) sieht im Fernsehen zum Beispiel *„... eine Prothese für die häusliche Dialogschwäche"* (zitiert nach Livenet o.J.) und die französische Schriftstellerin Francoise Sagan (o.J.) schlussfolgert, dass das Fernsehen *„... aus dem Kreis der Familie einen Halbkreis"* (zitiert nach Zeit online 2009) mache.

Die Aufzählung der Zitate verdeutlicht die zahlreichen Vorurteile gegenüber dem Fernsehen, die letztlich in dem Beitrag von Ursula von der Leyen (2010) polemisch zusammengefasst werden, so dass eine reflektierte Auseinandersetzung mit dem Medium selten in den Blickpunkt der Öffentlichkeit gerät. Doch es gibt durchaus auch positive Thesen und Meinungen, die an dieser Stelle nicht verschwiegen werden sollen. Zum Beispiel reagiert Gregor Gysi (1993), Fraktionsvorsitzender der Linkspartei im Deutschen Bundestag, auf die Frage nach Gewalt und Sex im Fernsehen folgendermaßen:

> *„Sollen nackte Körper und Gewalt aus dem Fernsehen verbannt werden? Nackte Körper gehören glücklicherweise, Gewalt unglücklicherweise zu unserem Leben. Also lautet die eigentliche Frage: Soll das Leben aus unseren Fernsehern verbannt werden?"* (zitiert nach Höltgen 2008)

Mit dieser Aussage unterstreicht Gysi bereits 1993, dass die Inhalte des Fernsehens seiner Meinung nach die Gesellschaft abbilden und es einer Zensur gleichkäme, dies zu untersagen.

Auch der prominente TV-Moderator und Journalist Günther Jauch (o.J.) versteht die Problematik nicht. Er ist der Ansicht, dass jeder frei entscheiden kann, ob er fernsieht und falls ja, welche Sendungen konsumiert werden. Folgt man Jauch (o.J.), ist dementsprechend jeder für sich selbst verantwortlich. Der bloße Bildschirm strahlt keine bösartigen Kräfte aus. Mit den Worten *„Es gibt nichts Demokratischeres als einen Fernsehapparat: Man kann einschalten, umschalten und ausschalten"* (zitiert nach Melzer o.J.) bekräftigt er dies. Die Formulierung erscheint insofern allerdings etwas problematisch, als dass die Bürger einer Demokratie hier mit passiven Konsumenten verglichen werden, die bei Nichtgefallen letztlich nur resignieren können, indem sie den Fernseher wieder ausschalten.

Eine weitere, das Fernsehen verteidigende Meinung stammt von Gert Müntefering (2012). In einem Interview mit dem Studenten Ömer Akyün (2012) gibt der Mitentwickler der Kindersendung *Die Sendung mit der Maus* zu Protokoll:

> *„Es gibt kein Fernsehen, das nicht bildet. Die Frage ist, was verstehen wir unter Bildung und wie definieren wir dann öffentliche Beiträge. Ich verstehe unter Bildung eine offene, nicht an Institutionen gebundene Lebensaufgabe und Lebensmöglichkeit. Sie gedeiht freilich nicht passiv, sondern sie setzt kognitives und ästhetisches Werkzeug voraus - und somit die Möglichkeit zur Auswahl und Beurteilung. Sie ist nicht mit einer geglückten Berufswahl beendet, und die kulturellen Bereiche überschneiden sich."* und ergänzt *„Ich denke, dass das Fernsehen den Gesetzen der Massenkommunikation gehorcht und dass es gleichzeitig eine Projektionsfläche für die eigene Person mit Gelerntem und Erträumten darstellt"* (zitiert nach Akyün 2012, S. 54).

Müntefering (2012) vertritt folglich die Position, dass das Fernsehen zur Bildung von Kindern und Jugendlichen beitragen kann, aber in anderer Form, als die Schule dies tut. Lesen, Schreiben und Rechnen lernen die Kinder nicht am Fernsehen und auch die Lebensertüchtigung zum Beruf und Staatsbürger kann das Fernsehen nicht übernehmen. Nichtsdestotrotz besitzt das Fernsehen das Potenzial, bei den Schülerinnen und Schülern das Interesse für die Welterkundung zu wecken.

> *„Wenn Bildung Arbeit, gehobenes Bürgertum, Spezialist für was auch immer bedeutet, also zweckgerichtete Anstrengung, dann wird es schwierig und überschreitet die Grenze zu Schule, Bildungseinrichtung und nachhaltigem Lernen. Wenn die Welterfahrung und Welterkundung eine Verbindung eingehen kann mit spannendem und lustvollem Spiel,*

> *dann ist auch das Fernsehen gefragt - inclusive des Zugangs (auch über das Internet) zu den Archiven der Welt"* (zitiert nach AKYÜN 2012, S. 55).

Eine vergleichbare Sichtweise stammt von dem Familien- und Kommunikationsberater JAN-UWE ROGGE (2005). Er ist der Meinung, dass Kinder durch primär wissensvermittelnde Fernsehsendungen etwas lernen können, denn insbesondere Vorschulsendungen sind häufig so aufgebaut, dass sie an *„kindlichen Fähigkeiten und am kindlichen Wissen anknüpfen"* (ROGGE 2005, S. 100). Lerneffekte sind vor allem dann möglich, wenn das im Fernsehen vermittelte Wissen mit der direkten Erfahrung in Zusammenhang gebracht wird. Beispielsweise indem die Gegenstände, die in den Fernsehsendungen von Bedeutung sind, auch in der Wirklichkeit vorliegen, so dass die Kinder eigene Erfahrungen machen können. Ob es sich dabei allerdings um die gängige Praxis oder einen realitätsfernen Wunsch handelt, bleibt fraglich. Dementsprechend erkennt ROGGE (2005, S. 100) auch einen fundamentalen Nachteil der Wissensförderung durch das Fernsehen, denn das visuelle Medium vermittelt häufig eher ein *„Wissen über als ein Wissen um einen Sachverhalt"*, obwohl vor allem Vor- und Grundschulkinder über das Greifen lernen und begreifen.

Da der Medientheoretiker MARSHALL MCLUHAN (1968) davon überzeugt ist, *„... dass ein neues Medium irgendeines Zeitabschnitts von jenen, die sich die Schemata früherer Medien irgendwelcher Art zu eigen gemacht haben, schon bald als Pseudo eingestuft wird"* (zitiert nach JOHNSON 2006, S. 29), überrascht es nicht, dass das Fernsehen auch in anderen Medien (Film, Buch, Zeitung, Musik, ...) teils scharf kritisiert wird. Daher soll an dieser Stelle kurz auf zu diesem Zweck ausgewählte Musikstücke und Filme eingegangen werden.

Spätestens als die Gruppe THE BAGGLES 1979 feststellte, dass das Fernsehen auch auf die Musikbranche einen großen Einfluss nehmen würde (*Video killed the radio star*), nahmen kritische Stimmen von Musikern gegenüber dem *neuen* Konkurrenzmedium zu. Zu nennen ist sicherlich in erster Linie das Musikstück *57 Channels (And Nothin' On)* des 20-fachen Grammy-Gewinners BRUCE SPRINGSTEEN von 1992. In dem Lied wird vor allem kritisiert, dass trotz der großen Programmvielfalt, repräsentiert durch die zahlreichen (57) Fernsehsender, keine Sendung ausgestrahlt wird, die lohnend genug ist, um geschaut zu werden. Indirekt wird hier dementsprechend zumindest angedeutet, dass das Fernsehen zur Verdummung der Bevölkerung beiträgt. Konkret singt die Rocklegende SPRINGSTEEN von zahlreichen, teils energischen, aber allesamt vergeblichen Versuchen, eine ansprechende Sendung zu finden. Zunächst probiert er es beispielsweise mit einem Kabelanschluss (*„Man came by to hook up my cable TV. We settled in for the night my baby and me. We switched 'round and 'round 'till half-past dawn. There was fifty-seven channels and nothin' on."* (SPRINGSTEEN 1992)), bevor er später auf einen Empfang per Satellitenschüssel wechselt (*„Well no home entertainment was my baby's wish. So I hopped into town für a satellite dish."* (ebd.)). Am Ende bleibt ihm, trotz aller Mühe, nur die Möglichkeit mit einer 44er-Magnum das Problem aus der Welt zu räumen, indem er sich auf ziemlich amerikanische Art und Weise, treu dem Motto „Ich werfe den Fernseher aus dem Fenster", frustriert dem Apparat entledigt (*„So I bought a 44magnum it was solid steel cast. And in the blessed name of Elvis well I just let it blast. 'Til my TV lay in pieces there at my feet."* (ebd.)).

Mit dem Format der Talk-Shows, welches in den 1990ern und zu Beginn der 2000er Jahre seine Hochzeit erlebte, aber aktuell in vielen Fällen gegen inhaltsähnliche Pseudo-Reality-Sendungen ausgetauscht wird, setzen sich die Alternative-Rocker von Incubus in ihrem Lied *Talk Shows on Mute* kritisch auseinander. Der Songtext beschreibt, wie sinnlos und bildungsarm Talk-Shows sind und dass den Menschen dies häufig erst bewusst wird, wenn man die entsprechenden Sendungen ohne Ton verfolgt. Sänger Brandon Boyd beschreibt seine schockierende, ihn zu dem Text inspirierende Vision des Fernsehens folgendermaßen:

> *„I was on an airplane when a talk show began playing on the TVs. I decided to start narrating for the people, which is a really great game if you're ever bored enough. I realized a time will probably come when television will watch us if we're watching it, if that hasn't already happened, figuratively or literally. It sounded like some sort of pseudo-Big Brother nightmare, so I wrote it down"* (Boyd 2004a).

Boyd präsentiert damit eine hochinteressante Sichtweise auf die Mechanismen und die manipulativen Kräfte der Bildschirmmedien und stellt bedeutende und komplexe gesellschaftliche Fragen: Beobachten die Zuschauer tatsächlich das Geschehen auf dem Bildschirm oder werden die Zuschauer selbst mittlerweile so genau von den Verantwortlichen der Fernsehindustrie analysiert, so dass das Fernsehen seine Konsumenten, ohne dass diese es merken würden, manipuliert und die Menschen in der Ausübung ihres freien Willens und damit in der Bildung einer eigenständigen Identität gewissermaßen eingeschränkt werden? Ist dasjenige Individuum, welches kein Fernsehen schaut und von den auf dem Bildschirm dargebotenen, teils primitiven, Leitbildern abweicht, beispielsweise ein hochbegabter Naturwissenschaftler, nicht direkt ein Aussätziger, der in einer durch das Fernsehen eben nur eingeschränkt toleranten Gesellschaft gar nicht existieren kann? Diktiert das Fernsehen in der heutigen Zeit, welche Ziele im Leben erreicht werden müssen, dass die Ausbildung und die Ausübung eines Berufs nur der Mittel zum Zweck sind, um im Wohlstand lebend, einem Leben nachgehen zu können, dass von dem Bedürfnis nach Spaß gesteuert wird?

Brandon Boyd scheint diese Fragen mit einem klaren *Ja* zu beantworten. Nicht umsonst ist der erste Vers des Refrains (*„Come one come all into nineteen-eighty-four"* (Boyd 2004b)) eine allzu offensichtliche Anspielung auf den Roman *1984* von George Orwell. In dem Klassiker der modernen Weltliteratur zeichnet Orwell ein visionäres Schreckensbild eines totalitären Überwachungsstaats, indem die Menschenrechte rigoros eingeschränkt werden. Die Bevölkerung Britanniens wird durch die Elite des Staates in ständiger Angst gehalten. An der

Spitze dieses totalitären Staatssystems steht ein fiktiver Führer, der „Big Brother". Mit Hilfe der Televisioren, es handelt sich dabei um Fernseher mit eingebauter Kamera, die nicht abgeschaltet werden können, wird neben der Verbreitung der Parteipropaganda zudem eine totale Überwachung ermöglicht (vgl. KAISER 2002). In dieser sozial- und politkritischen Parabel über totalitäre Gesellschaftsverhältnisse scheint BOYD eine Metapher für seine Vision des Fernsehens gefunden zu haben. Unter diesem Gesichtspunkt verwundert es nicht, wenn er beispielsweise darüber singt, dass das Studiopublikum bei Talk-Shows für seine Aufmerksamkeit bezahlt wird (*„Pay an audience to care, impress me personality"* (BOYD 2004b)), dass sich Talk-Show-Teilnehmer wie Schafe verhalten, die aus Gründen der Konformität und aus Angst vor Demütigung immer genau das tun, was andere machen oder erwarten (*„The electric sheep are dreaming up your face"* (ebd.)) oder dass die Fernsehzuschauer, ohne es zu merken, wie Motten von den Sendungen angezogen werden (*„Make all forget that they're the moth"* (ebd.)). Damit überträgt der INCUBUS-Sänger ORWALLS Warnung vor der uneingeschränkten Vereinnahmung der Menschen durch eine Parteielite ziemlich drastisch auf die Wirkungskräfte des Fernsehens, so dass das Lied zu mehr als nur einer bloßen Talk-Show-Kritik verkommt.

Während die Fernsehkritik vor allem in *Talk Shows on Mute* fast schon philosophische Züge annimmt, wird im deutschsprachigen Raum vor allem kritisiert, dass durch die Fülle an Sendungen und Informationen der Mensch in seiner Wahrnehmung abstumpft und weder durch besonders erfreuliche noch durch besonders schockierende Nachrichten, auch nicht im realen Leben, überrascht wird. Ein Beispiel für einen solchen Beitrag zur allgemeinen Fernsehkritik ist *Fernsehen* (1990) von der deutschen Punk-Rock-Institution DIE TOTEN HOSEN. In dem Text heißt es unter anderem:

> *„Ehen vor Gericht, die Krimistunde, ne Euroshow und ne Raterunde, ein Erdbeben und einen Bier-Werbespot, einen Fernsehkoch und einen Hungertod, Ölpest, Tanzfest, Feuerschlucker, das Wetter von morgen, ein Bericht über Zucker, Mondlandung, Tipps für Garten und Haus, Attentatsopfer und die Sendung mit der Maus. Das hab ich alles schon mal gesehen. Sie haben's im Fernsehen gebracht"* (FREGE 1990).

Durch die Vermischung und Gleichsetzung völlig verschiedener Programminhalte wird zudem pointiert kritisiert, dass die Bedeutung der einzelnen Sendungen und Reportagen im Zeitalter des *Zapping* scheinbar verloren geht.

In einem weiteren gesellschaftskritischen Lied (*Hier kommt Alex)* beklagt TOTEN-HOSEN-Sänger ANDREAS FREGE (CAMPINO) zudem, dass das Fernsehen in einer von normativen Erwartungen geprägten Welt, in der abweichendes Verhalten nicht geduldet und häufig durch radikale Interventionen sogar forciert wird, zur einzigen Aufregung in einem vorgeschriebenen und festgelegten Alltag wird (*„In einer Welt in der man nur noch lebt, damit man täglich roboten geht, ist die größte Aufregung, die es noch gibt, das allabendliche Fernsehbild*" (FRE-

GE 1988)). Gerade im Zusammenhang mit dem zuvor zitierten Musikstück, indem dem Fernsehprogramm praktisch der gesellschaftliche Nutzen abgesprochen wird, wird die Medien- und Gesellschaftskritik besonders deutlich: Obwohl das Fernsehen die Menschen abstumpft und mit zahlreichen, meist bedeutungslosen Inhalten bombardiert, ist es dennoch die einzige Abwechslung im menschlichen Leben, eine durchaus erschreckende These.

Doch nicht nur im musikalischen Sektor wird das Fernsehen auf vielfältige Weise kritisiert. Auch im Medium Film wird das visuelle Konkurrenzmedium starker Kritik ausgesetzt und das, obwohl sich Filme selbst seit Beginn der Kino-Ära mit teils heftigen Anfeindungen auseinandersetzen müssen.

In DAVID FINCHERS umstrittenem Kultthriller *Fight Club* aus dem Jahre 1999, basierend auf der Buchvorlage von CHUCK PALAHNIUK (1996), wird mit Hilfe einer Geschichte über einen Schizophrenen (EDWARD NORTON, BRAD PITT), der sich von der inhumanen Industriegesellschaft abgestoßen fühlt und nach Schmerz verlangt, um sich wieder lebendig fühlen zu können, die zentrale Frage *Was kommt nach der Desillusion*? gestellt. Im Film heißt es dazu beispielsweise in einer zentralen Szene *„Wir wurden durch das Fernsehen aufgezogen in dem Glauben, dass wir alle irgendwann mal Millionäre werden, Filmgötter, Rockstars ... werden wir aber nicht!"* (PALAHNUIK 1996). Der Film wirft dem Fernsehen an dieser Stelle folgerichtig vor, bei den Zuschauern falsche Vorstellungen über das reale Leben und mögliche Perspektiven zu suggerieren. Betrachtet man unter anderem die Flut an Casting-Sendungen (*Deutschland sucht den Superstar, Das Supertalent, Popstars, The Voice, ...)* im Fernsehprogramm, erscheint diese Kritik nicht besonders abwegig. Gerade Kinder und Jugendliche sind hier besonders gefährdet, schließlich wird dem Fernsehen in dem vorliegenden Filmzitat auch vorgeworfen, die Menschen *aufzuziehen.*

Während sich *Fight Club* nur an einigen Stellen mit den Wirkungsmechanismen des Fernsehens auseinandersetzt, legt *Die Truman Show* von PETER WEIR den Fokus vollständig auf die kritische Betrachtung des Fernsehprogramms. Bereits 1998, also zu einem Zeitpunkt, lange bevor die zahlreiche Reality-Shows wie *Big Brother* das TV-Programm tatsächlich dominierten, monierte der australische Regisseur in seinem satirischen Film, dass die Fernsehmacher selbst den banalsten Alltag im Fernsehen ausschlachten würden. Zu diesem Zweck schildert der Film die Geschichte von TRUMAN BURBANK (JIM CARREY), der nur scheinbar ein völlig normales Leben auf der Insel Seahaven führt. In Wahrheit befindet sich diese Insel in einem gigantischen Fernsehstudio, ausgestattet mit unzähligen versteckten Kameras, die jede Bewegung von TRUMAN aufzeichnen und in die ganze Welt übertragen. Der Protagonist selbst ist dabei die einzige Person, die sich dessen nicht bewusst ist. Sowohl seine Adoptiveltern *(Truman wurde von der Produktionsfirma als Baby adoptiert!),* sein bester Freund, seine Frau und alle weiteren Menschen sind bezahlte Hollywood-Schauspieler, die gesteuert von dem rigorosen Regisseur CHRISTOF nur Rollen spielen. CHRISTOF ist der Ansicht, dass seine künstlich geschaffene Fernsehwelt ein perfektes Gegenbild zur verkommenen Realität ist und die Zuschauer ausgerechnet von dem unwissenden und getäuschten TRUMAN lernen sollen, was Moral und Aufrichtigkeit ist. Als eines Tages, die Serie läuft zu diesem Zeitpunkt seit 10.909 Tagen rund um die Uhr, ein Scheinwerfer vom Himmel bzw. der Decke des Fernsehstudios fällt, kommt TRUMAN seinem eigenen Drama langsam auf die Schliche. Regisseur WEIR verfolgt mit seinem Film mehrere zentrale Anliegen: Er kritisiert gleichermaßen die Manipulation durch die Medien als auch den Konformismus und die Kommerzialisierung des Fernsehens. Gerade der Aspekt der Kommerzialisierung wird in dem Film auf ironische Art und Weise dargestellt. Als TRUMANS Frau in einer Szene ein Getränk serviert, hält sie das Produkt beispielsweise in eine versteckte Kamera und wirbt werbespotähnlich: *„Ich mache dir jetzt eine Tasse von dem neuen Mococo-Drink. Nur natürliche Kakaobohnen von den oberen Hängen des Mount Nicaragua ohne künstliche Süßstoffe"* (NICCOL 1998). Neben der Medienkritik wird zudem der Voyeurismus der Fernsehzuschauer gesellschaftskritisch analysiert. So wird in einer Szene erwähnt, dass viele Zuschauer den Fernseher sogar nachts laufen lassen würden, um TRUMAN beim Schlafen zuzusehen, während sich ein anderer Zuschauer darüber beklagt, dass man nie etwas sieht, wenn TRUMAN duscht, weil die Kameras immer nur den Vorhang zeigen würden. Deutlicher als in *Die Truman Show* wurde Fernsehkritik in einem Film wohl nie wieder dargestellt. Bedenkt man, dass seit 1998 die Anzahl solcher Reality-Soaps im realen TV drastisch zugenommen hat, erscheint diese Kritik durchaus visionär gewesen zu sein. Als weitere fernsehkritische Filme können *Tötet Smoochy* (2002) und *Pleasantville* (1998) genannt werden.

3.2 Interview mit Eva Radlicki (ZDF)

Da die meisten vorgestellten und in der Öffentlichkeit wahrgenommen Meinungen dem Fernsehen vorwerfen, vor allem negative Auswirkungen auf die soziale, körperliche und schulische Entwicklung von Kindern und Jugendlichen zu haben, erscheint es sinnvoll, im Gegenzug auch die Fernsehmacher selbst zu Wort kommen zu lassen. In diesem Zusammenhang hat sich EVA RADLICKI, aktuell Redaktionsleiterin „Information in der Hauptredaktion Kinder und Jugend" des öffentlich-rechtlichen Fernsehsenders ZDF, stellvertretend bereit erklärt, dem Verfasser einige Fragen auf elektronischem Weg zu beantworten. Da sie vor ihrer Tätigkeit beim ZDF in pädagogischen Projekten direkt mit Kindern gearbeitet hat (vgl. ZDF o.J.), wird man ihr nur schwer unterstellen können, bei der Gestaltung des Fernsehprogramms, soweit es ihr möglich ist, nicht am Wohle der Kinder interessiert zu sein. Zwar bestätigt sie, dass Kinder auch aus Gründen der Unterhaltung und Langeweile fernsehen, bekräftigt aber gleichzeitig, dass die Bildungsangebote des ZDF umfassend wahrgenommen werden und erläutert, dass es deutliche Unterschiede in der Programmgestaltung zwischen den öffentlich-rechtlichen und den privaten Sendern gibt. Selbst dem plakativen Vorwurf, dass das Fernsehen lediglich Fantasielosigkeit und Bewegungsmangel verbreiten würde, tritt Radlicki entschlossen, aber immer argumentativ, entgegen und unterstreicht damit, dass eine einseitige, von Vorurteilen geprägte Debatte über die Einflüsse des Fernsehens auf die kindliche Entwicklung weder dem Medium, noch der Arbeit *einiger* Fernsehsender gerecht wird.

M. Fey: In meiner eigenen Studie habe ich unter anderem die Ursachen des kindlichen und jugendlichen Fernsehkonsums untersucht. Bildungs- und Wissensformate sowie anspruchsvolle Unterhaltung sind dabei aber eher die Ausnahme. Wie können Sie sich das erklären?

E. Radlicki: Zu den Motiven und Funktionen der Medien- und insbesondere der Fernsehnutzung der Kinder liegen seit vielen Jahren zahlreiche empirische Forschungsergebnisse vor, die belegen, wie vielfältig sie sind. Von Kindern am häufigsten genannt werden die Motive Unterhaltung und Beseitigung von Langeweile. Medien fungieren für Kinder auch zur Regulation von Stimmungen, und sie werden zur Bewältigung von Alltagserfahrungen oder aus eskapistischen Motiven genutzt. Im Medienensemble kommt dabei dem Fernsehen eine wichtige Rolle zu. Fernsehangebote helfen Kindern dabei, Hinweise zur Bewältigung entwicklungsbedingter Themen und aktueller Problemlagen zu finden, und sie liefern Anregung für die Ausformung ethisch-normativer Orientierungen und personaler Vorbilder. Daneben erwarten sie vom Fernsehen auch, dass ihr Wissen erweitert wird. Neben Spaß und Spannung ist das eines der wichtigsten Motive für Kinder fernzusehen. Sie sind zum Beispiel neugierig darauf, zu erfahren, wie etwas funktioniert (z.B. Technik), was in der Welt passiert (z.B. Nachrichten) oder wie es sich mit bestimmten sozialen Phänomenen verhält (z.B. Ehescheidung). Auch möchten sie Einblick in für sie unbekannte Lebensbereiche (z.B. Berufswelt) oder generell in die Erwachsenenwelt bekommen.

Das ZDF strahlt sowohl im Rahmen seines Kinderprogramms am Wochenende als auch im KiKA Wissensformate für Kinder aus. Sie werden im KiKA zur Hauptsehzeit der Kinder gesendet und stoßen dort auf große Akzeptanz bei den Kindern. Noch höher fällt die Akzeptanz bei dem monothematischen Wissenmagazin *pur+* aus, das ebenfalls am Abend im KiKA ausgestrahlt wird.

M. Fey: Haben Bildungs- und Wissensformate im Fernsehen überhaupt eine Chance genügend Zuschauer zu finden? Was müsste sich vielleicht ändern?

E. Radlicki: Von daher haben wir nicht nur die Chance „genügend" Zuschauer zu finden – nein – die jungen Zuschauer finden und schätzen unser Angebot.

Der Bildungsauftrag des öffentlich-rechtlichen Fernsehens steht neben dem der Information, der Beratung, der Unterhaltung und insbesondere der Kultur. Bildungssendungen gehören daher zum Kernangebot des öffentlich-rechtlichen Fernsehens. Sie haben sich dem Wettbewerb zu stellen, der in Deutschland durch eine breite und vielfältige Angebotspalette an Free TV-Programmen einer der schärfsten weltweit ist. Er stellt stetig neue Qualitäts- und Attraktivitätsanforderungen an die Programme, und gerade Bildungsprogramme unterliegen besonderen Anstrengungen, um attraktiv zu sein. Seit Anbeginn zeichneten sich die Bildungsprogramme von ZDF und ARD dadurch aus, dass sie attraktive Vermittlungsformen anstrebten.

Das ZDF ist mit *Faszination Erde, Abenteuer Wissen, Abenteuer Forschung* und *Terra X* als populäre Wissens- und Wissenschaftssendungen 2011 erfolgreich. Zu den meistgesehenen Sendungen dieses Genres im deutschen Fernsehen gehören die ZDF-Formate *Faszination Erde* mit deutlich mehr Zuschauern als bei dem von ProSieben ausgestrahlten Wissensmagazin *Galileo*.

M. Fey: Hat das Fernsehen Ihrer Meinung nach noch einen Bildungsauftrag und falls ja, wie sieht der aus? Unterscheiden sich hier private und öffentlich-rechtliche Sender?

E. Radlicki: Natürlich haben wir diesen Auftrag – er ist im Staatsvertrag verankert. Und wir kommen diesem Auftrag auch in vielfältiger Form nach. Das ZDF-Kinderprogramm bietet mehr und höherwertige Wissensformate als jeder andere deutsche Sender. Mit unserem Kinderprogramm zeigen wir, dass öffentlich-rechtliche Informationskompetenz bereits im Kinderprogramm beginnt.

Als öffentlich-rechtliches Kinderfernsehen bedient das ZDF mit seinen Informationsprogrammen das Bedürfnis von Kindern nach altersgemäßer Information, Bildung und auch Unterhaltung. Kinder haben ein grundlegendes Bedürfnis, neues Wissen zu erwerben und sich mit diesen erworbenen Bausteinen ein Bild davon zu schaffen, wie die Welt funktioniert. Ein öffentlich-rechtliches Informationsprogramm für Kinder muss bildende, beratende und emotional berührende Programmteile enthalten und den spezifischen Bedürfnissen von Kindern Rechnung tragen.

Kinder erwarten ein Programm, das ihnen Spaß macht und in dem sie etwas Neues erfahren. Unsere Informationsprogramme sind Sendungen, die einen hohen Gebrauchswert für Kinder haben. Die Themen in unseren Programmen kommen aus dem Alltag der Kinder, orientieren sich an der kindlichen Neugier und die Gestaltung ist zielgruppengerecht. Unsere Wissensformate sind nicht die Fortsetzung des Schulunterrichts mit anderen Mitteln. Wohl aber wollen wir Kinder durch unsere Formate auf die Anforderungen der Zukunft vorbereiten. Die Informationsprogramme leisten einen Beitrag bei der Vermittlung von Werten, mit unseren Programmen liefern wir Denkanstöße und vermitteln Sozialkompetenz. Wir bieten Orientierung, die Kinder für ihr Familienleben, für die Schule, im Umgang mit der Natur und für ihre Entwicklung von Werten brauchen.

Mit der Markenkompetenz des ZDF Kinderinformationsprogramms trägt das ZDF profilbildend zum Angebot des Kinderkanals bei. Die (Regel)Sendungen der Redaktion Information genießen höchste Reputation und werden immer wieder – jenseits der guten Einschaltquoten und den ausgesprochen positiven Rückmeldungen seitens der Zuschauer – auch von Fachjurys mit Preisen für gutes Kinderfernsehen ausgezeichnet.

pur+ gibt es seit über 15 Jahren im ZDF tivi und natürlich auch im Ki.Ka. In der Sendung werden Wissenschaft, soziale Themen, Life-Style, Sport und emotionale Geschichten gleichermaßen behandelt. Eric Mayer ist der Reporter und Moderator der Sendung und probiert alles aus! Und es gibt in jeder Sendung Experimente - auch zusammen mit Kindern und Familien. Wer pur+ schaut, der weiß, wie sich Fakten anfühlen – und behält sie auch.

Löwenzahn
Das ‚Löwenzahn-Prinzip' ist die einzigartige Kombination aus spannenden und witzigen fiktionalen Geschichten, erklärenden Animationen und bildstarken Dokumentationen. Löwenzahn vermittelt seit über 30 Jahren Erkenntnisse und weckt die Entdeckerlust der jungen Zuschauer. Die Auseinandersetzung mit den Dingen soll bei ihnen im besten Falle (umwelt)bewusstes Denken und Handeln nach sich ziehen. Fritz Fuchs (Guido Hammesfahr) ist der Held der Sendereihe Löwenzahn. Er lebt mit Hund Keks in einem Bauwagen in Bärstadt. Von hier aus starten seine kleinen oder größeren Ermittlungen, weil er Fragen zu allem hat, was um ihn herum passiert. In jeder Folge entdeckt und erlebt er Neues und vermittelt dadurch ganz nebenbei das Wissen um das jeweilige Phänomen. Mit großer Neugier, aber niemals kritiklos, stellt er sich auch schwierigen Fragen aus Natur, Umwelt und Technik.

logo!
Kinder müssen auch mit politischen Krisensituationen in angemessener Weise konfrontiert werden. Unsere Informationsprogramme – und im Falle der aktuellen Berichterstattung insbesondere logo! - spiegeln Realität wieder, ordnen sie ein, machen sie verständlich und liefern Angebote zum Umgang mit den Ereignissen. Unsere Kindernachrichtensendung liefert dem Zuschauer täglich Antworten, auch zu ausgesprochen abstrakten Themen, über Zugänge, die sich an ihren Erfahrungen, ihren Werten und Emotionen orientieren.

Terra MaX
Lässt Geschichte lebendig werden, bedient die kindliche Fantasie und vermittelt im Rahmen erzählter Geschichten anspruchsvolle historische Fakten.

Vergleichbare Sendungen finden sich bei den Privatsendern gar nicht oder nur bedingt. Darin liegt ein deutlicher Unterschied.

M. Fey: Ganz allgemein: Glauben Sie, dass Medien wie das Fernsehen dazu beitragen, Bildung zu vermitteln oder eher dazu beitragen Bewegungsmangel und Fantasielosigkeit zu fördern?

E. Radlicki: Ja, wir vermitteln Bildung.

Alle unsere Wissensformate basieren auf einem modernen Konzept des Lernens. Wir geben keine fertigen Lösungen vor. So kann z.B. Fritz Fuchs in Löwenzahn lustvoll scheitern und zeigt damit, dass dies auch ein Schritt zur Erkenntnis sein kann. Wir regen unsere Zuschauer bei der Konstruktion ihrer Wissensstände an und fördern damit eigene Aktivität und eigenes Denken.

Nach dem traditionellen Konzept des Lernens, der Instruktion, soll das, was im Kopf der Wissenden/der Redaktion ist, in die Köpfe der Zuschauer. Dieser veraltete Ansatz produziert ein eher träges Wissen, das meist wertlos ist, da der junge Zuschauer nicht weiß, wie er es konkret anwenden soll. In unseren Sendungen arbeiten wir mit (konstruktivistischen) Ansätzen, die von der Eigenkonstruktion von Wissen durch die Kinder ausgehen. In den ZDF Kinderinformationsprogrammen geht es um die Vermittlung von Kenntnissen, die es Kindern ermöglichen, eigene Erkenntnisse hervorzubringen. Dadurch regen wir Kinder an, eine eigenständige Weltsicht zu entwickeln. Und natürlich auch aktiv zu werden.

An Bewegungsmangel und Fantasielosigkeit trägt niemals nur ein Faktor – also z.B. das Fernsehen - schuld. Beides entsteht durch ein familiäres und soziales Umfeld, in dem wenig Wert auf körperliche und intellektuelle Aktivität gelegt wird.

Um in einem Vergleich zu sprechen: nicht die leckere Schokolade ist daran schuld, wenn jemand deutlich übergewichtig ist, sondern ein generell falscher Umgang mit Lebensmitteln und dem eigenen Körper.

M. Fey: Vielen Dank!

3.3 Manfred Spitzer vs. Steven Johnson: zwei extreme Meinungen

Nach diesem einleitenden Überblick über prominente Meinungen zum Thema „Fernsehen und Bildung” sollen in diesem Abschnitt die Ansichten der zwei einflussreichsten Autoren, Manfred Spitzer (2012) und Steven Johnson (2006), gegeneinandergestellt werden. Beide Autoren verteidigen ihre radikalen Thesen teils populistisch, besonders interessant wird dies allerdings dadurch, dass dies vor dem Deckmantel der Wissenschaft und einer augenscheinlich handfesten Argumentation geschieht. MANFRED SPITZER (2012, S. 12) behauptet in *Vorsicht Bildschirm* beispielsweise besorgniserregend, dass es aufgrund der Bildschirmmedien, wie dem Fernsehen, 2020 in *„Deutschland jährlich etwa 40.000 Todesfälle durch Herzinfarkt, Gehirninfarkt, Lungenkrebs und Diabetes-Spätfolgen geben"* wird; hinzu kämen *„jährlich einige hundert zusätzliche Morde, einige tausend zusätzliche Vergewaltigungen und einige zehntausende zusätzliche Gewaltdelikte gegen Personen"* und warnt daher explizit vor der Nutzung entsprechender Geräte. In *Everything Bad is Good for You* (dt. *Neue Intelligenz: Warum wir durch Computerspiele und TV klüger werden*) vertritt STEVEN JOHNSON (2006, S. 26) genau die entgegengesetzte Meinung. Er befürwortet die Unterhaltungsmedien, da die populäre Kultur in den letzten Jahrzehnten immer komplexer geworden sei und somit den Verstand, für den Zuschauer kaum wahrnehmbar, schärfen würde.

3.3.1 *Vorsicht Bildschirm* von Manfred Spitzer

Nachdem MANFRED SPITZER, Neuropsychologe der Universität Ulm, vor einigen Jahren bereits von strombetriebenen Unterrichts- und Lernmedien abriet, warnt er seit 2006 nun explizit vor jeglichen Bildschirmen und meint damit u. a. die Nutzung von Computerspielen, Internet und Fernsehen. Vor allem mit Hilfe von Untersuchungen aus der Gehirnforschung und daraus konstruierten Kausalketten führt SPITZER negative Einflüsse des Fernsehens auf nahezu alle Bereiche der kindlichen Entwicklung an. Selten fließen dabei die tatsächlich auf dem Bildschirm präsentierten Inhalte in seine vorgelegten Studien und Argumente ein. In den meisten Fällen genügt dem Mediziner bereits ein Blick auf die Fernsehdauer, um seine Schlussfolgerungen zu ziehen und seine radikale Meinung zu untermauern.

Zunächst warnt MANFRED SPITZER (2012, S. 13 f) vor gesundheitlichen Risiken, die durch einen überhöhten Fernsehkonsum eintreten können. Der *mögliche* Zusammenhang zwischen Fernsehdauer und Gesundheit ist demnach zwar noch ein relativ junger Untersuchungsgegenstand, doch haben seit 1984 bereits über 50 Studien nachweisen können, dass der Konsum von Bildschirmmedien dosisabhängig zu Übergewicht führt. DIETZ und GORTMAKER (1985) gelten hier mit ihrem Forschungsbericht *Do we fatten our children at the television set?* als Vorreiter. Die Studie konnte im Querschnitt belegen, dass derjenige, der dick ist, viel fernsieht und im Längsschnitt zeigen, dass derjenige, der als Kind viel ferngesehen hat, auch einige Jahre später als Jugendlicher dick ist (vgl. SPITZER 2012, S. 20 f). Besonderen Wert legt der Autor allerdings auf eine von ROBERT HANCOX und seinen Mitarbeitern durchgeführte Längsschnittstudie, die im Jahre 2004 veröffentlicht wurde, da sie Auskunft über die langfristigen Konsequenzen des Fernsehkonsums in der Kindheit für den Gesundheitszustand von Erwachsenen gibt. Die entsprechenden Daten wurden im neuseeländischen Dunedin gesammelt. In den Jahren 1973/74 wurden dazu zunächst alle Kinder, die vor Ort geboren wurden, erfasst. Ab einem Alter von drei Jahren fanden dann in regelmäßigen Abständen von zwei bis drei Jahren Befragungen der Familien bzw. später der Probanden selbst bezüglich des durchschnittlichen Fernsehkonsums an Wochentagen sowie Wochenenden statt. Am Ende der Studie nach 26 Jahren war es den Forschern immer noch möglich, 96 % der Probanden hinsichtlich ihres Gesundheitszustands zu befragen. Dafür wurden u. a. Körpergröße, Gewicht, Blutdruck und Belastungswerte bestimmt und zusätzlich durch eine Fragebogenaktion der sozioökonomische Status der Herkunftsfamilie erfasst (vgl. SPITZER 2012, S. 29). Die Ergebnisse, die SPITZER (2012, S. 30) in seinem Buch präsentiert, sind erstaunlich:

> *„Je länger die Kinder in der Kindheit und Jugend (also im Alter zwischen fünf und 15 Jahren) vor dem Fernseher saßen, desto größer war ihr BMI. Man konnte weiterhin berechnen, dass 17 % des Übergewichts der Erwachsenen auf das Konto des Fernsehkonsums in der Kindheit gingen".*

Zudem beschreibt SPITZER (2012, S. 33), dass dieser Zusammenhang selbst dann bestehen bleibt, wenn der sozioökonomische Status der Herkunftsfamilien aus den gewonnen Daten ausgefiltert wird, so dass der Mediziner folgert: *„Es sind weder die dicken Eltern noch das schon vorbestehende eigene Übergewicht als Kleinkind für den Zusammenhang [...] verantwortlich"* (ebd., S. 33). Vielmehr fördern drei konkret benennbare Mechanismen diesen negativen Einfluss des Fernsehens auf die gesundheitliche Verfassung der Zuschauer, wobei SPITZER (2012, S. 33 ff) auf LUDWIG und GORDMARKER (2004) verweist: Zum einen kann in dem vor dem Fernseher verbrachten Zeitraum keiner körperlichen Aktivität nachgegangen werden. Die Amerikaner bezeichnen seit jeher Menschen, die sehr viel Zeit vor dem Fernseher verbringen, nicht umsonst als *Couch-Potato*. Zum anderen fördert der Fernsehkonsum schlechte Essgewohnheiten, da die Schauspieler im Fernsehen beispielsweise eher ungesund essen. Außerdem verbraucht derjenige, der viel Zeit vor dem Bildschirm verbringt, weniger Energie, was zur Ansammlung dieser Energie in Form von Fettgewebe und damit zu Übergewicht führt. Das Risiko dieses Übergewichts liegt vor allem in möglichen Folgeerkrankungen wie Bluthochdruck und schlechten Cholesterinwerten.

MANFRED SPITZER warnt aber nicht nur vor fernsehbedingtem Übergewicht. Er kritisiert auch, dass das Fernsehen die Erfahrung und die Aufmerksamkeit negativ beeinträchtigt, und setzt sich in diesem Zusammenhang genauer mit dem menschlichen Gehirn auseinander. Eine der wesentlichsten Funktionen des menschlichen Gehirns besteht darin, Regelmäßigkeiten in der Erfahrung der Umgebung zu entdecken bzw. zu repräsentieren. Besonders wichtig ist, dass das Gehirn derartige Erfahrungen in sich aufnimmt, um diese anschließend für das zukünftige Handeln in der Welt zu verwenden (vgl. SPITZER 2012, S. 4 f). Hier sieht SPITZER (2012, S. 5) bereits die Gefahr negativer Einflüsse des Fernsehens, da er davon überzeugt ist, dass die Bildschirmmedien diese notwendigen Erfahrungen der Kinder manipulieren, so dass sich Körper und Geist des Kindes verformen. Die Auswirkungen dieser Verformung sind auf die gesamte Lebensdauer hin gültig, da die Erfahrungen, die in den frühen Lebensjahren gewonnen werden, sich im Laufe der Zeit immer weiter manifestieren und sich nur schwer korrigieren lassen. Dies erklärt bereits, warum Kinder und Jugendliche am stärksten von den Risiken der bildschirmhaften Medienlandschaft betroffen sind. Sie haben noch den längsten Lebensweg vor sich. SPITZER (2012, S. 6) erklärt dies so:

> *„Die vom Bildschirm vermittelten Wahrnehmungserlebnisse sind von der normalen Wahrnehmung verschieden, und es ist dieser Unterschied in der Form der Wahrnehmungserlebnisse, der sich auf die Formung des kindlichen Geistes ungünstig auswirkt".*

Dazu liefert der Autor die passende Studie der Kaiser Family Foundation (RIDEOUT; VANDEWATER; WARTELLA 2003): In den USA verbringen bereits Zweijährige im Durchschnitt mehr als zwei Stunden, das entspricht 13 bis 22 % ihrer Erfahrungen im wachen Leben, vor den verschiedenen Bildschirmen, so dass sie um die Möglichkeit gebracht werden, eigene Erfahrungen im Umgang mit der Welt zu sammeln. Dies ist problematisch, denn wir Menschen nehmen die Welt eigentlich nicht passiv wahr, sondern filtern aktiv das für uns Wesentliche aus den vielen uns erreichenden Stimuli aus und verarbeiten dies weiter. Die Fähigkeit, bestimmte Stimuli zu bevorzugen, zu behandeln und in dessen Folge ihre Wahrnehmung überhaupt erst zu ermöglichen, nennt SPITZER (2012, S. 67) die *selektive Aufmerksamkeit.* Zudem bezieht sich SPITZER (2012, S. 83) in diesem Zusammenhang auf die *Konzentrationsfähigkeit*, die für die bessere Verarbeitung bestimmter Reize sorgen soll. Dies ist aber nur möglich, wenn aufgrund der richtigen Erfahrungen mit der Welt in der Kindheit innere Strukturen vorhanden sind, da sonst keine bewusste Steuerung der Konzentration möglich ist. Nun ist der Kritiker aber davon überzeugt, dass das Fernsehen eben nicht in der Lage ist, eine aktive Wahrnehmung der Umwelt zu ermöglichen, so dass sowohl die *selektive Aufmerksamkeit* als auch die *Konzentrationsfähigkeit* der Kinder durch das Fernsehen beeinträchtigt wird:

> *„Bildschirme können noch so bunt sein, das Bild ist flach und der Inhalt verglichen mit der Wirklichkeit arm, riecht nicht, schmeckt nicht und lässt sich nicht anfassen. [...] Wenn kleine Kinder einen wesentlichen Teil ihrer Zeit vor dem Bildschirm verbringen, dann muss das ungünstige Auswirkungen auf deren Entwicklung haben"* (ebd., S. 6 f).

Hier sieht SPITZER (2012, S. 90) u. a. auch die Ursache dafür, warum der Inhalt des Fernsehprogramms keinerlei Bedeutung für seine kritische Haltung hat. Alleine die Form der gelieferten Erfahrungen ist schädlich, da das Kind die Welt noch nicht kennt und betrachtete Objekte und Szenen eben gerade nicht aus Vorerfahrungen ergänzen kann, so dass neben der Gesundheit auch die Aufmerksamkeit und die Erfahrung durch das Fernsehen negativ beeinträchtigt werden, was dauerhaft durchaus auch auf negative Auswirkungen hinsichtlich des Schulerfolgs schließen lässt.

In einem weiteren Kapitel stellt sich SPITZER (2012, S. 121) die für diese Arbeit wohl bedeutendste Frage: *„Gibt es einen Zusammenhang zwischen Fernsehkonsum und Schulerfolg?"*

Wenig überraschend beantwortet er auch diese *rhetorische* Frage umgehend mit einem klaren *Ja*. SPITZER (2012, S. 121) argumentiert vor allem auf der Basis zweier Studien. Dabei handelt es sich einerseits um eine Pionierarbeit des Teams um den Freiburger Psychologen MICHAEL MYRTEK (2003) und andererseits um die Untersuchung des Würzburger Psychologen MARCO ENNEMOSER (2003b).

Anhand von 223 Schülerinnen und Schülern der Altersgruppen 11 und 15 Jahre führten MYRTEK und seine Freiburger Forschungsgruppe 2003 eine Untersuchung durch, um die Wirkung des Fernsehens auf die Rezipienten in Bezug auf ihre Schulleistung zu ermitteln. Die Studie ist insofern interessant, als dass neben psychologischen auch die psychophysiologischen Variablen Puls und körperliche Bewegung erhoben wurden, da diese es ermöglichen, sowohl die körperliche Aktivität (viel Bewegung und hoher Puls) als auch die emotional-mentale Aktivität (keine Bewegung und hoher Puls) abzuleiten. Anhand dieser Daten können die körperliche, die geistige und die emotionale Beanspruchung von Kindern und Jugendlichen objektiv gemessen werden. Die kontinuierliche Aufzeichnung der Daten erfolgte über insgesamt 23 Stunden, so dass es den Forschern möglich war, eine sehr genaue Beschreibung des Alltags der Probanden anzugeben, welcher neben der Freizeit und der Schule natürlich auch Aspekte wie den Schlaf enthält. Um diese Daten mit den Fernsehgewohnheiten der Kinder in Bezug zu setzen, wurden die Probanden in die beiden Gruppen „Wenigseher" (11järige: 0,8 Stunden; 15jährige: 2,9 Stunden) und „Vielseher" (11jährige: 1,1 Stunden; 15jährige: 3,3 Stunden) eingeteilt, so dass die Unterschiede zwischen diesen beiden Gruppen auf das gesehene Programm und verschiedene Aktivitäten und Fähigkeiten bestimmt werden konnten (vgl. SPITZER 2012, S. 125 ff). Die vielfältigen Erkenntnisse bestätigen SPITZERS (2012, S. 125 ff) Vermutungen: Wer viel fernsieht, bewegt sich weniger, liest seltener, ist häufiger allein, führt seltener Gespräche und ist häufiger zu Hause. Diese Gegebenheiten münden gebündelt, wie unten abgebildet, letztlich in einer durchschnittlich schlechteren Deutschnote der Vielseher gegenüber den Wenigsehern (etwa -0,3).

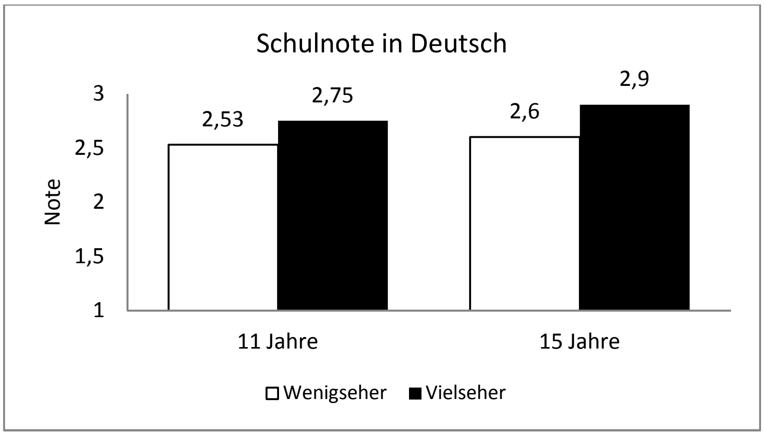


Abbildung 3.1 Schulnote in Deutsch nach Myrtek (Quelle: Spitzer (2012))

Myrtek (2003, S. 458) kommentiert das Ergebnis wie folgt:

> *„Zu nennen sind die schlechteren Schulleistungen der Vielseher, die sich vor allem in der Deutschnote niederschlagen. Vermutlich hängt dieser Befund mit der [...] geringeren Lesehäufigkeit der Vielseher, den seltener geführten Gesprächen und den geringeren Kontakten mit Freunden zusammen. Damit ist ein unmittelbarer Bezug zur PISA-Studie hergestellt. Kritisch ist zudem der [...] Bewegungsmangel der Vielseher zu bewerten. Es sind gerade jene Aktivitäten vermindert, die einen hohen Energieaufwand erfordern, z.B. das Radfahren und zu Fuß gehen. [...] Damit wird späteren Erkrankungen des Kreislaufs, des Stoffwechsels und der Gelenke Vorschub geleistet".*

Die zweite große Untersuchung, die Spitzer (2012, S. 133 ff) in dem Themenfeld Schulerfolg vorstellt, ist die oben angekündigte Studie der Würzburger Arbeitsgruppe, die den Fernsehkonsum im Kindergartenalter und die Lesefähigkeit in der Schule in Bezug gestellt hat. Anhand der insgesamt 332 teilnehmenden Kinder aus zwei Altersgruppen (letztes Kindergartenjahr und zweites Schuljahr) wurden unter anderem Daten über das Medienverhalten und die schriftsprachlichen Kompetenzen der Kinder gewonnen. Im Unterschied zu den meisten anderen Studien wurde die Mediennutzung sehr genau und detailliert über ein eigens entwickeltes Tagebuch erhoben, indem u. a. genau eingetragen werden sollte, in welchem Zeitraum die Medien genutzt wurden. Zudem konnten die Kinder Angaben zu dem Programm oder der Anwesenheit der Eltern beim Fernsehen machen. Anhand der erhobenen Daten konnten die Kinder in die drei Gruppen „Wenigseher" (0,25 Stunden), „Normalseher" (1 Stunde) und „Vielseher" (2 Stunden) eingeteilt werden. Die sprachlichen und schriftsprachlichen Kompetenzen wurden über standardisierte Testverfahren ermittelt.

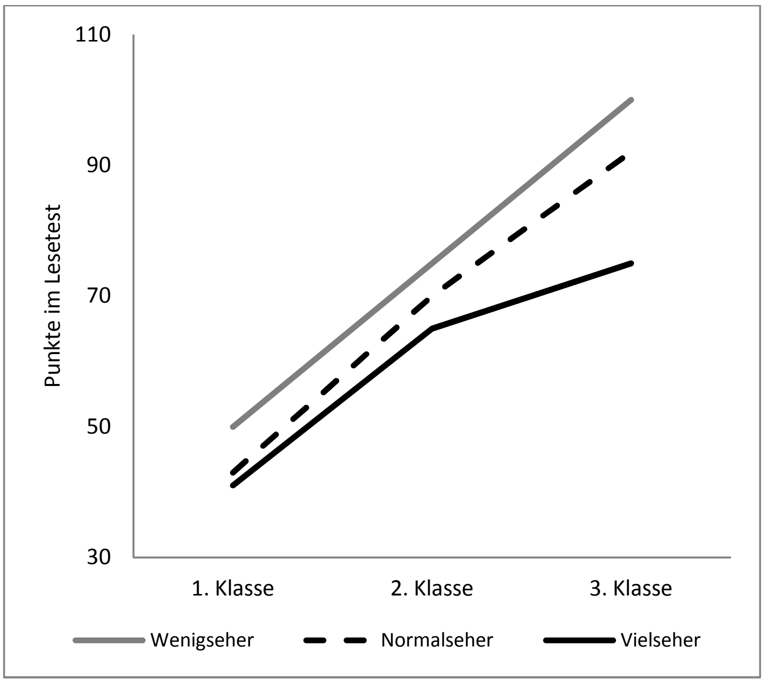


Abbildung 3.2 Punkte der Kinder im Lesetest nach Ennemoser (Quelle: Spitzer (2012))

Das obige Diagramm präsentiert das Ergebnis der Untersuchung: die Entwicklung der Leseleistung von der ersten bis zu dritten Klasse für die Kinder der jüngeren Untersuchungsgruppe in Abhängigkeit vom Fernsehkonsum. Zwar verbessern sich alle Kinder kontinuierlich, doch zeigt sich, dass besonders die Vielseher im Verlauf der zweiten und dritten Klasse deutlich hinter den Leistungen der weniger fernsehenden Mitschüler zurückblieben. Vielseher sind also nicht nur schlechter im Lesen, sie lernen es auch langsamer als Wenigseher. Auf diese Daten stützt Spitzer (2012, S. 134) seine Schlussfolgerung, dass sich bereits der Fernsehkonsum im Kindergarten sehr deutlich auf das Lesen lernen in der Grundschule auswirkt. Auch die Analyse der älteren Probanden unterstreicht das Ergebnis.

Ein weiterer gravierender Befund dieser Studie ist, dass das Vielsehen auf weniger intelligente Kinder einen wesentlich problematischeren Einfluss besitzt als auf intelligente Kinder. Die Studie zieht diese Erkenntnis aus einem Wortschatztest: Während sich bei intelligenten Kindern die Wortschatzentwicklung recht unabhängig vom Fernsehkonsum gleichermaßen positiv entwickelt, fallen minder intelligente Vielseher deutlich ab (vgl. ebd., S. 134 ff). Da es nicht seiner Theorie entsprechend ist, verschweigt Spitzer aber leider, dass im Gegensatz dazu Normalseher mit geringem IQ sogar Vorteile gegenüber Wenigsehern mit geringem IQ aus ihrem Fernsehkonsum ziehen können, wie die nachfolgenden Grafiken, die die erreichten Punkte bei diesen Wortschatztests entsprechend des jeweiligen IQ zeigen, verdeutlichen. Das regelmäßige Fernsehen von etwa einer Stunde pro Tag hat hinsichtlich des Wortschatzes dementsprechend sogar einen positiven Einfluss auf die Kinder mit niedrigem Intelligenzquotienten.

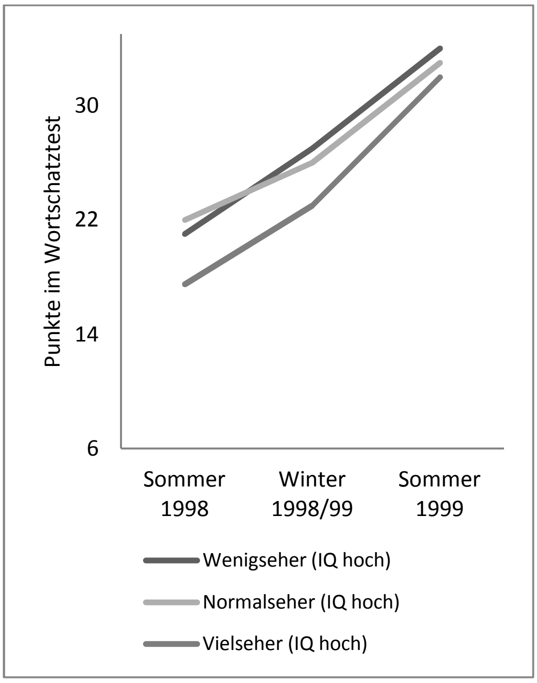


Abbildung 3.3 Wortschatz bei hohem IQ nach Ennemoser (Quelle: Spitzer (2012))

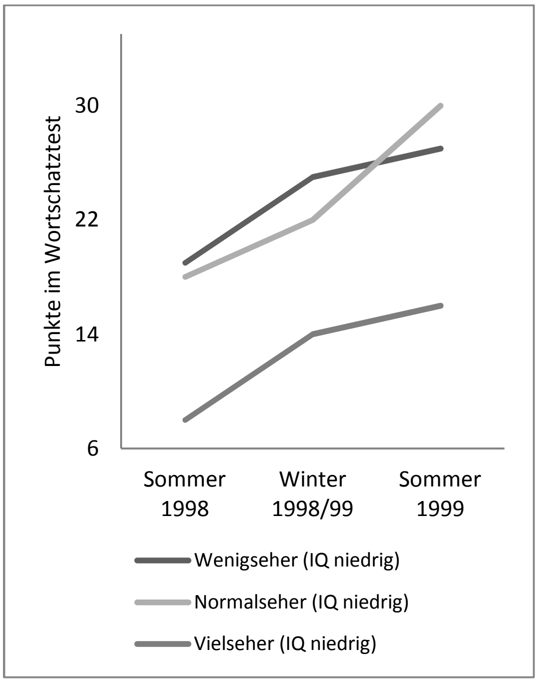


Abbildung 3.4 Wortschatz bei geringem IQ nach Ennemoser (Quelle: Spitzer (2012))

Auf Basis dieser beiden Untersuchungen sieht Manfred Spitzer den Schulerfolg im Zuge hohen Medienkonsums stark gefährdet. Leider bieten diese Daten allerdings keinen direkt nachweisbaren Rückschluss auf Kompetenzen außerhalb des schriftsprachlichen Bereichs, wie etwa den mathematischen Fähigkeiten, so dass die gewonnenen Erkenntnisse nicht zwingend Aufschluss über einen negativen Effekt des Fernsehens auf den gesamten Schulerfolg geben müssen. Daher ist Spitzers (2012, S. 154) Schlussfolgerung, Fernsehen mache dumm, mit Vorsicht zu genießen.

Im weiteren Verlauf seines Kreuzzugs gegen die Bildschirmmedien bezieht sich Spitzer (2012, S. 93 ff und 155 ff) zudem noch auf die *möglichen* negativen Einflüsse des Fernsehens auf die Gewalt- und Konsumbereitschaft der jungen Zielgruppe. Zudem vergleicht er die Bildschirmmedien mit der Umweltverschmutzung, da die Gesellschaft irgendwann einen ähnlich hohen Preis aufgrund der negativen Auswirkungen zu zahlen habe (vgl. ebd., S. 246). Schlussendlich zieht Manfred Spitzer (2012, S. 281) das reißerische Fazit:

> *„Bildschirmmedien machen dick und krank, wirken sich in der Schule ungünstig auf die Aufmerksamkeit und das Lesen lernen der Kinder aus und führen zu vermehrter Gewaltbereitschaft sowie tatsächlicher Gewalt".*

3.3.2 *Everything Bad is Good for You* von Steven Johnson

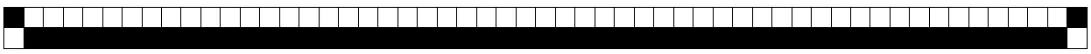
Steven Johnson dagegen ist wohl der bekannteste Verfechter der Bildschirmmedien. Das Besondere an seinen Ausführungen in *Everything Bad is Good for You* ist dabei, dass er weder mit Hilfe empirischer Studien, noch inhaltlich argumentiert. Ganz im Gegenteil: Er argumentiert strukturell. Zu Beginn seines Beitrags über das Fernsehen gesteht Johnson (2006, S. 74 f) zwar ein, dass das Fernsehen im Vergleich zu den ständig Entscheidungen fordernden Computerspielen relativ passiv sei und bestreitet auch nicht, dass Schimpfwörter, Gewalt und nackte Brüste im täglichen Fernsehprogramm zugenommen haben, doch ist er gleichzeitig davon überzeugt, dass die geistigen, emotionalen und kognitiven Herausforderungen an Mensch und Gehirn mindestens in gleichem Maß zugenommen haben. Um diese Ansicht belegen zu können, vergleicht er das Fernsehprogramm bzw. die wichtigsten Programmformate vergangener Jahre mit den heutigen Sendungen.

Laut Johnson (2006, S. 75) gibt es unterschiedliche Arten der Passivität beim Fernsehen. Natürlich gibt es Sendungen, bei denen sich die Zuschauer nur auf das Sofa setzen und gedankenverloren in den Fernseher starren. Folgt man dem amerikanischen Autor, wird gerade diese Form der Passivität durch das Fernsehen in seiner heutigen Form aber immer seltener. Immer häufiger treten dagegen Serien in Erscheinung, bei denen die Zuschauer komplexe Zusammenhänge erkennen müssen. Dabei geht es u. a. um das Erfassen verschiedener Erzählfäden oder das Füllen von Lücken basierend auf Vorwissen, Fantasie, Meinungsaustausch oder eigenen Recherchen. Somit ähneln sich Fernsehserien und Bücher laut Johnson (2006, S. 76) mehr denn je, da beide Medien Aufmerksamkeit, Geduld, Beharrlichkeit oder die Analyse von Erzählstrukturen erfordern.

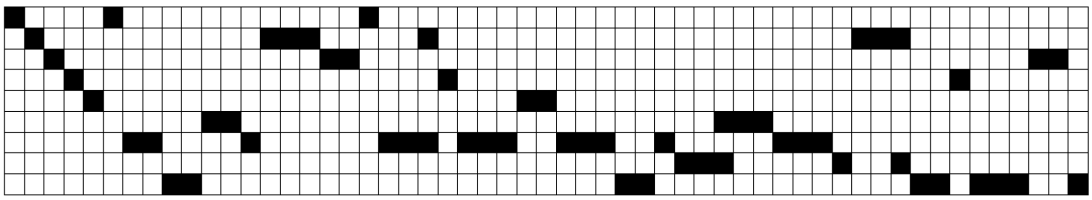
Steven Johnson (2006, S. 77) benutzt den Begriff des *Multithreading* als Anhaltspunkt für die Komplexität einer Fernsehserie. Vereinfacht ausgedrückt bedeutet *Multithreading* das Verknüpfen verschiedener Handlungsfäden. Gerade in modernen Dramaserien wie *Lost, Die Sopranos, West Wing* oder *24* tritt dieses Phänomen verstärkt auf, während ein Zuschauer in alten Serien wie *Polizeibericht* meist nur einem einzigen Handlungsfaden zu folgen hatte. Laut Johnson (2006, S. 77) ist dies eine geistige Herausforderung, die ein rundum geistig passiver Zuschauer nicht leisten kann. Um zu verdeutlichen, wie viele verschiedene und komplexe Erzählstrukturen heutige Fernsehsendungen aufweisen, präsentiert Johnson (2006, S. 81) zu den chronologisch nacheinander produzierten Serien *Polizeibericht, Starsky & Hutch, Polizeirevier Hill Street* und *Die Sopranos* seine sogenannte Schläferkurven:

Polizeibericht (jede Episode)

Starsky & Hutch (jede Episode)



Polizeirevier Hill Street (Episode 85)



Die Sopranos (Episode 8)

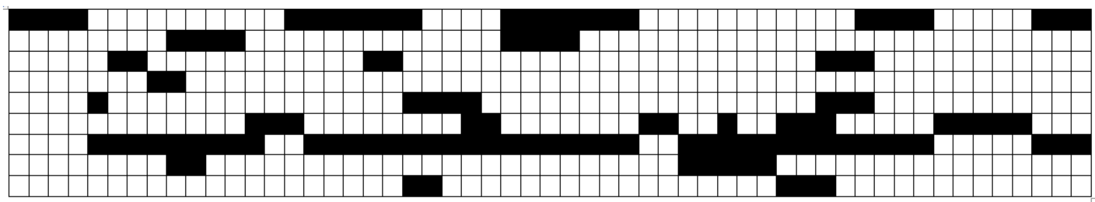


Abbildung 3.5 Schläferkurven (Quelle: Johnson (2006))

Die Diagramme belegen, dass die modernen Fernsehdramen die meisten Handlungsfäden besitzen und diese miteinander immer stärker verbinden. Der *Polizeibericht* (ab 1951) folgte seinerzeit noch einer einzigen stringenten Handlung, während bereits bei *Starsky & Hutch* (ab 1975) zu Beginn und Ende jeder Episode eine weitere Erzählebene in Form einer Rahmen-handlung eingeführt wurde. *Polizeirevier Hill Street* (ab 1981) revolutionierte das Genre des Fernsehdramas mit der Erweiterung auf bis zu 9 Erzählfäden, was anfangs noch zur Überforderung der Zuschauer führte. Dieses Konzept wurde im Laufe der Jahre immer weiter ausgebaut. So besitzt die achte Episode von *Die Sopranos* (ab 1999) zwar ebenfalls neun Erzählfäden, diese überschneiden sich aber wesentlich häufiger und werden untereinander auf intelligente und komplexe Art und Weise verknüpft. Die Einzelhandlungen in *Polizeirevier Hill Street* liefen dagegen noch nahezu unabhängig voneinander ab und trafen nur in entscheidenden Schlüsselszenen aufeinander. JOHNSON (2006, S. 81) ist sich sicher, dass dieses moderne Muster intelligenzfördernd ist und die Zuschauer die vielschichtigen Serien so begeistert aufnehmen, weil sie über Jahre langsam an diese geistige Herausforderung herangeführt wurden. Ob es sich bei den angegebenen Sendungen allerdings um kindgerechtes sowie repräsentatives Fernsehen handelt, darf zumindest stark bezweifelt werden. Nichtsdestotrotz wird *Multithreading* als wichtigste strukturelle Eigenschaft des modernen und komplexen Fernsehdramas beschrieben (vgl. ebd., S. 83).

Ein zweiter Bereich des modernen Fernsehens, der nach JOHNSON (2006, S. 84) dazu führt, dass die potenziellen Zuschauer geistig und kognitiv aktiv sein müssen, um einer Serie folgen zu können ist ein Bruch mit einer traditionellen Film-Konvention. Dabei handelt es sich um das sogenannte *Blinkende-Pfeil-Phänomen*: Der *blinkende Pfeil* ist definiert als ein *offensichtlicher* Hinweis, wie etwas funktioniert, damit ein Zuschauer der Handlung folgen kann, ohne denken zu müssen. Typische Beispiele für solche Signale sind eine düstere und bedrohliche musikalische Untermalung von Szenen, in denen der Bösewicht auftaucht, der Bauplan einer Bank kurz vor dem geplanten Banküberfall oder die ausführliche Beschreibung elektronischer, physikalischer oder chemischer Geräte. Die Anzahl dieser *blinkenden Pfeile* ist im Laufe der letzten Jahre, glaubt man JOHNSON (2006, S. 84), immer weiter zurückgegangen. Heutige Serien, als hervorragendes Beispiel dürfte hier die Mysteryserie *Lost* dienen, haben oft Geheimnisse und Rätsel in nahezu jeder Szene, über die die Zuschauer sich mindestens bis zur nächsten Episode den Kopf zerbrechen können. Beispielsweise kann allein durch das bewusste Zurückhalten von Informationen dieser Effekt erreicht werden (vgl. ebd., S. 86). Der Unterhaltungsfaktor moderner Serien liegt dementsprechend nicht alleine im passiven Konsum, sondern vielmehr in der kognitiven Arbeit, Informationslücken zu füllen, eigene (fantasievolle) Theorien zu schmieden und fehlende Details selbst einzufügen (vgl. ebd., S. 88). Dies ist sicherlich ein Punkt, der auch auf viele Sendungen zutrifft, die von Kindern geschaut werden.

Noch kindbezogener wird dies aber, wenn JOHNSON (2006, S. 94 f) sich dem Genre der Sitcoms, hier repräsentiert durch *Scrubs* und *Die Simpsons,* zuwendet. In diesem Fall seien derartige externe Informationen nötig, um die Witze überhaupt verstehen zu können. Während dem Zuschauer also heutzutage selbst in Sitcoms grundlegende Informationen über die Vergangenheit der Personen präsent sein müssen und der Humor auf *„elegant verflochtenen Plots* *und obskuren Anspielungen und Verweisen"* (ebd., S. 94 f) basiert, beruhten die Witze früher auf *banalen* Gemeinheiten und *vorhersehbaren* Slapstickeinlagen. JOHNSON sieht dementsprechend auch in dieser Veränderung eine Quelle geistiger und kognitiver Bildung beim Zuschauer durch den Fernsehkonsum.

Der nachfolgende Grund, aus dem STEVEN JOHNSON (2006, S. 118) davon überzeugt ist, dass das Fernsehen den potenziellen Zuschauer klüger machen kann, ist letztlich eine Kombination der bisher genannten Argumente: Da die sozialen Netzwerke der an den Dramaserien beteiligten Charaktere einerseits immer komplexer werden und andererseits vorzugsweise auf blinkende Pfeile als Verständnishilfe verzichtet wird, entsteht eine weitere kognitive Herausforderung für das Fernsehpublikum. Es muss sich dieses Konstrukt merken, um Veränderungen in den sozialen Beziehungen verstehen, vorauszuahnen und analysieren zu können. JOHNSON (2006, S. 119 ff) versucht diesen Argumentationsstrang zu belegen, indem er das soziale Netzwerk der modernen Serie *24* mit dem sozialen Netzwerk der älteren, aber weltberühmten Serie *Dallas* (1978) vergleicht. In den nachfolgenden Beziehungsmodellen stehen die hellen Linien für Beziehungen, die dem Zuschauer klar sein müssen, um der Handlung folgen zu können. Die dunklen Linien zeigen zusätzlich diejenigen Beziehungen, die wichtige Ereignisse innerhalb der Geschichten auslösen.

Dallas

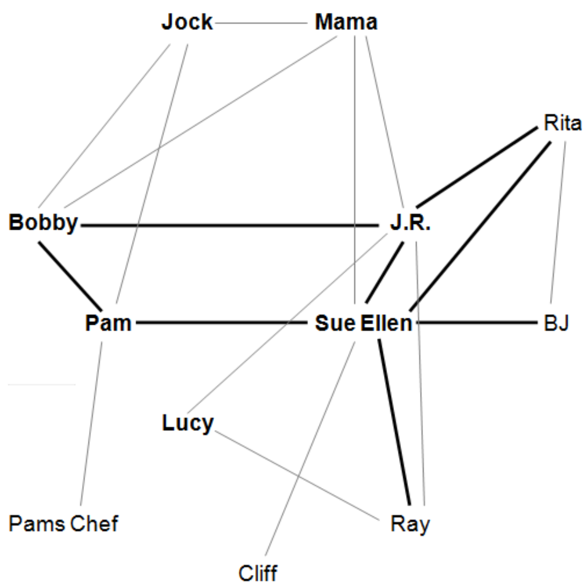


Abbildung 3.6 soziales Netzwerk Dallas (Quelle: Johnson (2006))

Das soziale Netzwerk aus *Dallas* konzentriert sich vornehmlich auf die verschiedenen Beziehungen einer einzigen Familie (EWING). So symbolisiert die helle Verbindung zwischen JOCK und Pam zum Beispiel, dass Jock es missbilligt, dass Pam arbeiten gehen und erst später Kinder bekommen möchte.

Für diejenigen, die dieses Beziehungskonstrukt bereits als komplex und kompliziert empfinden, fügt Johnson (2006, S. 121) nochmals hinzu, dass es in jeder Folge von *Dallas* genug deutliche Hinweise in Form blinkender Pfeile an den Zuschauer gibt, die das soziale Netzwerk immer wieder aufs Neue erklären.

Da *24* wesentlich mehr Handlungsstränge als *Dallas* besitzt und gleichzeitig eine Geschichte um drei Familien (Bauer, Palmer, Drazens), eine fiktive Anti-Terror-Einheit (CTU) inklusive zahlreicher Wendungen und Maulwürfe erzählt, ist es naheliegend, dass auch das soziale Netzwerk wesentlich komplexer ist. Dennoch verzichtet *24* nahezu vollständig auf blinkende Pfeile und fordert den Zuschauer damit kognitiv heraus (vgl. ebd., S. 121).

24

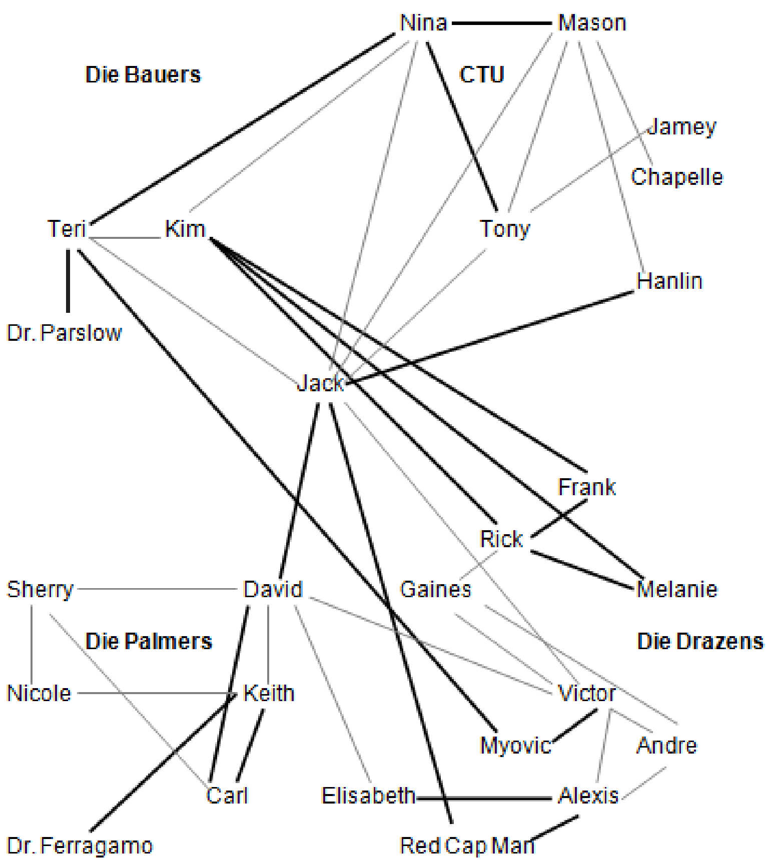


Abbildung 3.7 soziales Netzwerk 24 (Quelle: Johnson (2006))

Wie der Vergleich der beiden Abbildungen augenscheinlich belegt, hat das Genre der Dramaserien in den letzten dreißig Jahren auch im Sinne der gesellschaftlichen Komplexität einen sagenhaften Sprung gemacht, so dass Johnson (2006, S. 121) darauf besteht, dass sich *24* in diesem Punkt sogar auf Augenhöhe mit Kleinstadt- und Gutshof-Romanen von Jane Austen oder George Eliot befindet.

Glaubt man Manfred Spitzer, haben derartige Sendungen allerdings wenig mit dem tatsächlichen Fernsehkonsum von Kindern und Jugendlichen zu tun. Dem realen Fernsehprogramm kommt Johnson (2006, S. 100) dementsprechend erst richtig nahe, wenn er fragt, wie es sich mit dem „*Mist im Fernsehen*" verhalte. Folgt man dabei der strukturellen Argumentationslinie, dem Vergleich von früher und heute, kommt man scheinbar zu dem Ergebnis, dass sich selbst der *Mist*, zumindest im Vergleich zum *Mist* von früher, verbessert habe. Dieser *Mist*, umfasst heute verschiedene Formate des sogenannten *Reality-TV*.

Reality-TV ist ein Sammelbegriff für verschiedene Unterformate, die die Realität in verschiedenster, meist vereinfachter Form wiedergeben sollen. Hier kann beispielsweise zwischen Reality-Soaps (*The Osbourns*), Reality-Spielshows (*Big Brother, Deutschland sucht den Superstar, Der Bachelor*), Selbstverbesserungs- bzw. Hilfsshows (*Die Super Nanny, Rach: Der Restaurant-Tester, Raus aus den Schulden*) und Scripted-Reality (*Mitten im Leben, die Schulermittler*) unterschieden werden.

Die Frage, warum sich die Zuschauer solche *niveauärmeren* Sendungen anschauen, beantwortet Johnson (2006, S. 101) überraschend: Während sich die Zuschauer in der öffentlichen Wahrnehmung lediglich an der Demütigung der an den Sendungen teilnehmenden Personen ergötzen, ist Johnson der Meinung, dass der hohe Konsum solcher Produktionen vor allem ein Beweis dafür sei, wie stark unsere aktuelle Popkultur vom Spielerischen dominiert wird.

Der Autor geht sogar noch weiter und erklärt, warum selbst solche Sendungen die Zuschauer *klüger* machen, als es vergleichbare Sendungen in der Vergangenheit taten: Die Zuschauer versetzen sich in die Charaktere in der virtuellen Welt, analysieren die Spielregeln, setzen sich mit den gruppen- und psychodynamischen Prozessen auseinander und versuchen die Intelligenz und Fähigkeiten der Kandidatinnen und Kandidaten einzuschätzen. Zudem überlegen sich die Zuschauer eigene geeignete Strategien, wie sie selbst beispielsweise im Dschungel (*Ich bin ein Star, holt mich hier raus*) überleben könnten, oder die anderen Kandidatinnen auszustechen, um den begehrten Traummann für sich zu gewinnen (*Der Bachelor*). Laut Johnson (2006, S. 105 f) liegt in diesen Selbstprojektionen der Erfolg des *Reality-TV*. Das TV trainiert dementsprechend heutzutage selbst in seinen schlechten Momenten immerhin noch die emotionale Intelligenz der Zuschauer. Für den Autor ist dies der große Gewinn dieser eigentlich banalen Form der Fernsehunterhaltung. Das Fernsehen sei zudem als einziges Medium in der Lage, darzustellen, wie Menschen mit anderen Menschen umgehen (vgl. ebd., S. 109). Zur Untermauerung seiner These stellt Johnson (2006, S. 112) im Vergleich zwischen einer Sendung aus dem heutigen *Reality-TV* und der älteren Sendung *Glücksrad* fest, dass sich früher im Büro niemand mit seinen Kolleginnen und Kollegen über die Strategie der *Glücksrad*-Teilnehmer ausgetauscht hat, während heutzutage in Schulen, Universitäten und Büros durchaus über die Auswahl der Frauen in *Bauer sucht Frau* diskutiert wird. Leider verschweigt Johnson, wie sich das Realtiy-TV auf die Intelligenz der Zuschauer auswirkt, die solche Sendungen dauerhaft, in hohem Maße und ausschließlich konsumieren.

Zusammengefasst besteht Steven Johnsons Fernsehverteidigung aus zwei großen Bereichen: Die *guten* Fernsehserien werden komplexer (Multithreading, blinkende Pfeile) als früher, so dass sie mehr denn je zum Mitdenken auffordern und selbst das *schlechte* Reality-TV fördert die soziale und emotionale Intelligenz, so dass uns das Fernsehen, zumindest im Vergleich zu früher, klüger macht.

3.3.3 Kritik

Sowohl die Warnungen von Manfred Spitzer als auch die Ausführungen und insbesondere der Argumentationsstil von Steven Johnson werden sehr kontrovers diskutiert. Dies vermag kaum zu verwundern, sind die jeweiligen Ansichten wie bereits eingangs angekündigt, doch ziemlich einseitig und radikal. Während Bildungspolitikerin Annette Schavan (2012) Spitzers Buch beispielsweise als *„eindringliches Plädoyer gegen den hohen Bildschirmkonsum von Kindern und Jugendlichen"* (zitiert nach Klappentext von Spitzer 2012) lobt, kritisiert Dirk Frank (2005, S. 1) von der Goethe-Universität in Frankfurt, dass sich Statistiken, wie die Spitzers (man betrachte: ohne Bildschirm-Medien gäbe es in den USA 10.000 Morde und 70.000 Vergewaltigungen weniger (vgl. Spitzer 2012, S. 8)), wohl auch für Autos, Durchlauferhitzer und elektrische Dosenöffner aufstellen lassen würden.

> *„Anstatt Kontexte konkret zu benennen, in denen Bildschirmmedien in der Tat negative Wirkungen zeigen können, schreibt* Spitzer *schon der bloßen technischen Apparatur geradezu diabolische Kräfte zu"* (Frank 2005, S. 2),

kritisiert er zudem. Denn obwohl Spitzer sehr auf den Schein einer wissenschaftlichen Beweisführung fokussiert sei, bleibt seine Argumentation oberflächlich (vgl. ebd., S. 2).

Zu Steven Johnsons *Everything Bad ist Good for You* lassen sich ebenfalls zahlreiche positive Kritiken finden. Zum Beispiel schreibt Frank Rich (2006) von der New York Times, dass das Buch unentbehrlich für jeden sei, der diese Kultur verstehen will (zitiert nach Klappentext von Johnson 2006) und der Economist (2006) fügt hinzu, dass Johnson ein *„willkommenes Gegenmittel gegen den Pessimismus und das besorgte Händeringen all derer, die in der Popkultur lediglich Niedergang und Verderbnis sehen wollen"* (zitiert nach Klappentext von Johnson 2006) gelungen ist. Nichtsdestotrotz kritisiert Frank (2005, S. 4) nach Spitzer auch Johnson, da es sich ebenso um polemische Vereinfachungen handle, nur eben andersherum, da Johnson weder inhaltlich argumentiere, noch die Dauer des Fernsehkonsums betrachte, sondern sich lediglich mit der *Machart* der Sendungen auseinandersetze.

3.4 Exemplarische wissenschaftliche Studien

Wie sowohl die Gegenüberstellung der beiden populärsten Bücher als auch die populären Meinungen gezeigt haben, lassen sich Argumente und Schreckgespenste für beide Seiten der Medaille *Fernsehen* finden. Da die Wahrheit womöglich wie so oft in der Mitte liegt, erscheint es sinnvoll, an dieser Stelle den Bereich der subjektiven Meinungen von Einzelpersonen zu verlassen und sich auf möglichst objektive empirische Studien aus dem Themenfeld *Fernsehen und Bildung* zu beziehen, die beiden Positionen Gehör verschaffen.

Im Folgenden werden dazu drei besonders interessante Forschungsansätze präsentiert. Zunächst werden zwei Studien des Kriminologischen Forschungsinstituts Niedersachsens vorgestellt. Im Anschluss daran wird der ebenfalls fernsehkritische Mensch-Zeichen-Test beschrieben, welcher sich mit der bildhaften Wahrnehmung und dem Fernsehkonsum beschäftigt. Abschließend zeigt das Sesamstraßenexperiment, wie pädagogisch geprägte Fernsehsendungen die kindliche Entwicklung fördern können.

3.4.1 Kriminologisches Forschungsinstitut Niedersachsen

Die Kritiker der visuellen Medien berufen sich zumeist auf Studien des kriminologischen Forschungsinstituts Niedersachsens (KFN). Das im Jahre 1979 vom ehemaligen Justizminister Niedersachsens, Prof. Dr. Schwind, gegründete und aktuell von Christian Pfeiffer geleitete Forschungsinstitut gilt als unabhängige und interdisziplinär arbeitende Einrichtung, die praxisorientierte kriminologische Forschung fördern möchte (vgl. KFN o.J.a). Seit 1998 führt das KFN Schülerbefragungen zu wechselnden inhaltlichen Schwerpunkten durch, um von den polizeilichen Kriminalstatistiken unabhängige Informationen über das Ausmaß an Jugendgewalt, das als zentrales Thema des Instituts angesehen wird, zu erhalten.

3.4.1.1 Mediennutzung und Schulleistung

Im Jahre 2005 wurde der Schwerpunkt erstmals auf die Mediennutzung von Kindern und Jugendlichen gelegt. Das zentralste Projekt *Mediennutzung und Schulleistung* (2005 bis 2012) von Projektleiter Mössle verfolgt dabei die Idee, die Mediennutzung in der Freizeit und die schulische und gesellschaftliche Leistungsfähigkeit von Schülerinnen und Schülern miteinander in Bezug zu setzen (vgl. KFN o.J.b). Dazu sind neben den schulischen Leistungen der Probanden auch das familiäre und soziale Umfeld sowie das sonstige Freizeitverhalten analysiert worden, um über das Medienverhalten hinausgehende, die Schulleistung betreffende, Einflussfaktoren in die Ergebnisinterpretation miteinbeziehen zu können. Schließlich zeigte sich bei den Befunden der PISA-Studie, dass schlechte Schulleistungen u. a. mit dem Bildungsniveau der Eltern, dem sozioökonomischen Status, dem ethnischen Hintergrund, der Region und dem Geschlecht erklärt werden können (vgl. KLEIMANN et al. 2007, S. 4).

Im Rahmen dieses Projekts fand 2005 eine fragebogenbasierte Querschnittsbefragung von 5.529 Viertklässlern und 14.301 Neuntklässlern aus zehn verschiedenen westdeutschen Städten und Regionen zu ihrer Mediennutzung, ihren Familienverhältnissen, ihrem Freizeitverhalten, ihren Schulleistungen und ihrem Sozialverhalten statt (vgl. MÖSSLE et al. 2007, S. 48 ff). Die Lehrkräfte erhielten ebenfalls einen Fragebogen zum Ausfüllen, in dem neben zusätzlichen Informationen über jedes einzelne Kind, eine vorläufige Schulempfehlung für eine weiterführende Schule sowie Schulnoten in mehreren Hauptfächern notiert werden sollten (vgl. ebd., S. 49). Um eine bessere Vergleichbarkeit der Schulnoten herbeizuführen, wurden diese für die einzelnen Fächer am jeweiligen Klassenmittelwert standardisiert, wobei der Klassendurchschnitt einem Wert von null entspricht, so dass eine positive Abweichung nach oben für eine bessere und eine negative Abweichung nach unten für eine im Vergleich zum Klassenmittelwert schlechtere Leistungen steht (vgl. ebd., S. 91 f). In dieser repräsentativen Befragung wurden korrelative Beziehungen zwischen den befragten Aspekten festgestellt (vgl. KLEIMANN 2011, S. 20): *„Die intensive Nutzung von Filmen [...] und hierbei insbesondere die Nutzung gewaltbetonter, wenig entwicklungsförderlicher Angebote geht mit schlechteren Schulleistungen einher"* (vgl. MÖSSLE et al. 2007, S. 99).

Im Folgenden werden die wesentlichen Resultate aufgeführt:

<u>(i) Geräteausstattung</u>

KLEIMANN et al. (2007, S. 4 f) stellten fest, dass bereits 36 % aller Viertklässler ein eigenes Fernsehgerät in ihrem Zimmer besitzen. Die Geräteausstattung der Jungen ist mit 42 % deutlich höher als die der Mädchen (31 %) (vgl. MÖSSLE et al. 2007, S. 72). Migrantenkinder verfügen häufiger über ein Fernsehgerät als einheimisch deutsche Kinder. Darüber hinaus scheint es auch einen Zusammenhang zwischen dem Bildungsniveau der Eltern und dem Gerätebesitz ihrer Kinder zu geben. Diese Vermutung lässt sich durch die Tatsache bekräftigen, dass Kinder, deren Eltern einen sehr niedrigen Bildungsabschluss (z.B. Hauptschulabschluss) vorweisen, viermal so häufig einen eigenen Fernseher erhalten, als wenn die Eltern die allgemeine Hochschulreife erreicht haben. Zusammenfassend verfügen vor allem Jungen, Migrantenkinder und Kinder mit niedrigem sozioökonomischen Status häufiger über ein eigenes Fernsehgerät. Wenig überraschend nimmt der Gerätebesitz im Laufe der Zeit zu, so dass wesentlich mehr Neunt- als Viertklässler einen eigenen Fernseher besitzen.

Es zeigt sich bereits bei der Medienausstattung der Kinder und Jugendlichen ein negativer Zusammenhang zu den Schulleistungen. Nach MÖSSLE et al. (2007, S. 92) stehen zum Beispiel Kinder der vierten Klasse, die keinen eigenen Fernseher oder keine eigene Spielekonsole besitzen, in den Schulfächern Mathematik, Deutsch und Sachkunde zwischen 0,1 und 0,2 Notenpunkten besser da, als ihre Mitschüler. Im Gegensatz dazu liegen Kinder, deren Zimmer mit visuellen Mediengeräten ausgestattet sind, zwischen 0,1 und 0,3 Notenpunkten unterhalb des Klassendurchschnitts. Von der Problematik sind vor allem die Jungen betroffen. Diese Diskrepanz verwundert nicht, konnte doch gezeigt werden, dass Jungen häufiger mit medialen Geräten ausgestattet sind.

Der Umfrage zufolge erhalten Kinder ohne eigene Geräte zudem erheblich häufiger eine Gymnasialempfehlung von ihrer Grundschule, wohingegen nur jedes fünfte Kind mit eigenen Geräten eine solche Empfehlung erhält (vgl. KLEIMANN et al. 2007, S. 9).

(ii) Zeitlicher Fernsehkonsum

Die Auswertung der Studie zeigt, dass Viertklässler an normalen Schultagen durchschnittlich 101 Minuten ihrer Freizeit vor dem Fernseher verbringen, während es an Wochenenden oder Feiertagen sogar 153 Minuten werden. Viertklässler, die über ein eigenes Gerät verfügen, sehen durchschnittlich etwa eine Stunde länger fern als Viertklässler ohne eigenen Fernseher (vgl. ebd., S. 6). Des Weiteren konnte belegt werden, dass Jungen täglich 22 Minuten länger vor dem Fernseher sitzen als Mädchen. Auch Kinder aus bildungsfernen Elternhäusern schauen länger fern als ihre Vergleichsgruppe. Besonders stark betroffen sind Migrantenkinder (vgl. MÖSSLE et al. 2007, S. 57 ff).

Die Ergebnisse für die Neuntklässler machen zusätzlich deutlich, dass die tägliche Fernsehdauer mit steigendem Alter zunimmt. Für die Neuntklässler gewinnt neben dem Fernseher zusätzlich vor allem das Internet an medialer Bedeutung. Dadurch opfern diese Probanden den visuellen Medien insgesamt noch wesentlich mehr Zeit (vgl. KLEIMANN et al. 2007, S. 9).

Auch beim Vergleich der Gruppen unterschiedlicher Mediennutzungsdauer stellten MÖSSLE et al. (2007, S. 92 f) fest, dass Vielseher in allen Fächern schlechtere Leistungen zeigen als die Mittel- und die Wenigseher. Eine geschlechtsspezifische Betrachtungsweise zeigt zudem, dass Mädchen, die besonders lange fernsehen, vor allem in Mathematik unterhalb des Klassendurchschnitts liegen, während Jungen eher Probleme in Deutsch aufweisen.

(iii) Programminhalte

Betrachtet man die Auswertung der konsumierten Programminhalte, wird deutlich, dass sich der eigene Fernsehbesitz nicht nur auf die Fernsehdauer, sondern zusätzlich auch auf den Inhalt der Medien auswirkt. Beispielsweise schauen Kinder der vierten Klasse, die über ein eigenes Gerät verfügen, *„doppelt so häufig Filme, die erst ab 16 Jahren frei gegeben sind oder keine Jugendfreigabe erhalten haben"* (KLEIMANN et al. 2007, S. 9) als Kinder ohne eigenen Fernseher.

Insgesamt wird auch bei der Analyse der Programminhalte deutlich, dass sowohl das Geschlecht als auch der Bildungs- und Migrationshintergrund der Kinder und Jugendlichen Einfluss auf den Anspruch des Fernsehprogramms haben, so dass es nicht weiter verwundert, dass auch die konsumierten Programminhalte in Zusammenhang zu den erfassten Noten stehen. Demnach korrelieren besonders schwache Schulleistungen mit dem Konsum von gewalttätigen und jugendgefährdenden Fernsehsendungen und -filmen (vgl. ebd.).

Zwar legen die dargestellten Ergebnisse der Studie nahe, dass die Fernsehnutzung tatsächlich in direktem Zusammenhang zu den beobachteten Schulleistungen steht, doch hält MÖSSLE (2007, S. 92) fest, dass *„... eine erhöhte Ausstattung mit Mediengeräten im Kinderzimmer ..."* und die Mediennutzungszeiten *„... sowohl Charakteristikum großer Bildungsferne der Eltern als auch einer nachlässigeren Medienerziehung ist"*. Daher *„... lässt dieser Befund isoliert betrachtet noch keinen Rückschluss darauf zu, inwieweit die Geräteausstattung ..."* und die Mediennutzungszeit *„... als eigenständiges Phänomen schlechtere Schulleistungen bedingen kann"*.

MÖSSLE et al. (2007, S. 96) gehen dementsprechend davon aus, dass der Einfluss des Medienverhaltens auf die Schulleistungen der Kinder nicht nur von der Medienausstattung und der Mediennutzungszeit abhängt, sondern auch von Faktoren, wie dem Migrationshintergrund, dem innerfamiliären Klima, dem Geschlecht der Kinder und vor allem dem Bildungshintergrund der Eltern. Daher wurden derartige Faktoren bei den oben genannten Untersuchungsergebnissen berücksichtigt, um ein bereinigtes und damit repräsentatives Ergebnis zu erhalten. Aus diesem Grund wurden Daten deutscher männlicher Kinder ohne Migrationshintergrund ausgewertet, die in einem angenehmen Familienklima aufgewachsen sind und deren Eltern einen mittleren bis hohen Bildungsabschluss besitzen. Laut den Auswertungsergebnissen (vgl. ebd., S. 96) zeigt sich, dass die mit einem Fernseher (und einer Spielekonsole) ausgestatteten Jungen außer im Fach Deutsch durchschnittliche oder leicht überdurchschnittliche Noten erhalten, während ähnliche Schüler ohne Gerätebesitz, überdurchschnittliche Leistungen erbringen und in Deutsch 0,25, in Sachkunde 0,3 und in Mathematik 0,4 Notenpunkte über dem Klassendurchschnitt liegen. Wird die von den Kindern vor dem Fernseher verbrachte Zeit untersucht, so ist zu erkennen, dass die Gruppe der Jungen mit günstigen familiären Ausgangsbedingungen bei intensiver Nutzung nur noch durchschnittliche Schulleistungen erbringen, bei mittelmäßiger Nutzung deutlich überdurchschnittliche und bei geringer Nutzung die besten Schulleistungen erzielen.

3.4.1.2 Berliner Längsschnitt Medien

Das Projekt *Mediennutzung und Schulleistung* wurde zum gleichen Zeitpunkt von der Längsschnittstudie *Berliner Längsschnitt Medien* begleitet. Dabei wurden Berliner Grundschüler jährlich zu den gleichen Aspekten befragt, mit dem Ziel herauszufinden, ob sich die Mediennutzung auch langfristig auf die befragten Bereiche auswirkt. Vertraut man einer Zwischenbilanz dieser Teilstudie, scheinen sich die Ergebnisse der Befragung von 2005 zu bestätigen.

Bei der Berliner Längsschnittstudie wurde außerdem die mathematische Intelligenz der Schülerinnen und Schülern der dritten Klasse getestet. Obwohl aus den Intelligenztests hervorgeht, dass sich einheimisch deutsche und Migrantenkinder in ihrem mathematischen Leistungspotenzial nicht unterscheiden, schneiden Letztere in der vierten Klasse im Fach Mathematik um 0,4 Notenpunkte schlechter ab (vgl. Kleimann et al. 2007, S. 10).

Daraus lässt sich schließen, dass es weitere Ursachen für die schlechteren Leistungen in der Schule geben muss. Die Intelligenz ist ein Zusammenspiel mehrerer Faktoren, so dass das soziale Umfeld (Familie, Schule, Peer-Groups) die schulischen Leistungen beeinflussen kann. Die Medienwirkungsforschung vermutet, dass die Leistungsunterschiede *„Folge divergierender Mediennutzungsmuster sind"* (vgl. ebd., S. 10).

Das würde bedeuten, dass die Befunde in direktem Zusammenhang zu dem Medienkonsum stehen. Durch den erhöhten Konsum visueller Medien, wie dem Fernsehen, vernachlässigen die Kinder beispielsweise ihre Schulaufgaben und es kann zu Leistungseinbrüchen in der Schule kommen. Diese Effekte sind jedoch in allen Bildungsschichten erkennbar. Dies meint, dass auch Kinder mit protektiven Rahmenbedingungen davor nicht geschützt sind. Daher ist anzunehmen, dass die prägnanten Leistungsdefizite bei den sozial benachteiligten Kindern eher auf familiäre Verhältnisse und häusliche Gewalterfahrungen zurückzuführen sind (vgl. Mössle 2009, S. 24).

3.4.1.3 Zusammenfassung der Studien des KFN

Als Endergebnis der beiden Studien kann festgehalten werden, dass sich das Fernsehen negativ auf den Schulerfolg auswirken kann, vor allem dann, wenn es besonders häufig genutzt wird und problematische Inhalte bevorzugt werden, sowie der Zugang durch ein eigenes Gerät erleichtert wird. Da diese Bedingungen vor allem bei Jungen, Migranten und bildungsfernen Kindern anzutreffen sind, besteht logischerweise insbesondere bei ihnen die Gefahr schulischer Leistungsdefizite. Doch selbst die Forscher stellen in ihrem Fazit fest, dass die Schulleistung auch von vielen weiteren Einflussfaktoren abhängt. Nach wie vor sind für den Schulerfolg nämlich vor allem der Bildungshintergrund der Familie, die gewaltfreie Erziehung und das Familienklima von Bedeutung. Dennoch gilt es festzuhalten, dass diese Faktoren die Mediennutzung maßgeblich beeinflussen, da zum Beispiel die Eltern über den Gerätebesitz im Kinder- und Jugendzimmer entscheiden. Nicht die Medien selbst machen die Kinder dementsprechend *dumm*, sondern deren falsche, altersunangebrachte Nutzung erhöht das Risiko schulischer Misserfolge. So betont auch Ursula von der Leyen (2006), dass *„… kein noch so restriktives Gesetz …"* den Missbrauch der Bildschirmmedien verhindern kann, *„… wenn es im Elternhaus keine Medienkultur gibt"* (zitiert nach Welt online 2006).

3.4.2 Der Mensch-Zeichen-Test

Auf der Suche nach Risikofaktoren für die kognitive Entwicklung von Kindern hat sich der Kinder- und Jugendarzt Peter Winterstein aus Baden-Württemberg, unterstützt durch Robert J. Jungwirth, in Form eines sehr interessanten Experiments unter anderem auch mit dem Einfluss eines erhöhten Fernsehkonsums auseinandergesetzt. Inspiriert von Studien, die einen negativen Effekt des Fernsehens zumindest auf die Lese- und Rechtschreibfähigkeiten von Schülerinnen und Schülern erahnen lassen (vgl. Spitzer 2012), hat Winterstein (2006, S. 1) ein altersadäquates Verfahren entwickelt, um den Einfluss des Medienkonsums auf die bildliche Wahrnehmung von fünf- bis sechsjährigen Kindern zu erforschen, da er der visuellen Wahrnehmung neben der auditiven Wahrnehmung und der sprachlichen Entwicklung enorme Bedeutung im Bezug auf die schulische Leistungsfähigkeit zuschreibt.

Winterstein (2006, S. 1) erklärt zu Beginn seiner Ausführungen, dass die bildhafte Wahrnehmung *„sich auf der Basis vorhandenen Wissens und neuer Signale, die von einem Objekt ausgehen"* in drei Phasen entwickelt. Die erste Stufe umfasst die sogenannte Kritzelphase, in welcher die einjährigen Kleinkinder Freude an Bewegungen entwickeln und Erfahrungen der Selbstwirksamkeit sammeln. Im Verlauf der zweiten Phase lernen die Kinder Kreise, Bögen und Punkte darzustellen, während sie in der ersten Phase nur in der Lage waren Striche zu „kritzeln". Die dritte Stufe ist davon geprägt, dass die Kinder langsam die Fähigkeit erlangen, Menschenformen zeichnerisch darzustellen. Da diese Prozesse mit der allgemeinen kognitiven Entwicklung des Kindes korrelieren, verfolgt Winterstein (2006, S. 2) bei seiner Untersuchung das Ziel, die bildhafte Wahrnehmung in Abhängigkeit vom kindlichen Fernsehkonsum zu bestimmen. Auf Basis dieser Ergebnisse sei es dann eben auch möglich, die Auswirkungen des Fernsehens auf die kognitiven Fähigkeiten der Kinder abzuschätzen. Weil Kinder während des Fernsehkonsums einerseits eine schnelle Bildfolge verarbeiten müssen und andererseits während dieser Zeit keinen motorischen, sozialen und kreativen Aktivitäten nachgehen, stellt der Kinder- und Jugendarzt die These auf, dass das Fernsehen negative Auswirkungen auf die bildhafte Wahrnehmung haben muss.

Um diese These zu überprüfen, ließ Winterstein in Anlehnung an die oben erwähnte dritte Stufe knapp 1.900 Kinder passenden Alters mit Papier und Bleistift ausstatten und gab ihnen die zunächst simpel klingende Aufgabe, Menschen zu malen. Bei diesem *Mensch-Zeichen-Test*, wie Winterstein sein Experiment nennt, sind weder perfekte anatomische Darstellungen gefordert, noch sind Schönheit und künstlerischer Anspruch der Zeichnungen von vorrangiger Bedeutung. Vielmehr ist es wichtig, dass die gemalten Menschen halbwegs vollständig aussehen und nachvollziehbare Proportionen aufweisen, so dass es den Testern möglich gemacht wird, die entstandenen Figuren nach vorbereiteten Kriterien zu bewerten. Grundsätzlich werden solche *Mensch-Zeichen-Tests* zwar bereits seit 1926 eingesetzt, doch ist es Winterstein 2004 gelungen, 13 Bewertungskriterien aufzustellen, die für Vorschulkinder auch tatsächlich relevant sind (vgl. ebd., S. 2).

Bewertet werden der Kopf, das Kopfhaar, die Augen, die Nase, der Mund, die Ohren, der Hals, der Rumpf (plastisches Ausführen und Form), die Arme, die Finger, die Beine und die Füße der Figuren. Jedes Kind konnte für seine Zeichnung folgerichtig zwischen null und 13 Punkten bekommen. Im Mittel erreichten die Kinder 9,5 Punkte (vgl. ebd., S. 2).

Parallel zu dem *Mensch-Zeichen-Test* wurde die Fernsehdauer der teilnehmenden Kinder bestimmt, wobei der durchschnittliche tägliche Wert bei 62,2 Minuten liegt (vgl. ebd., S. 2).

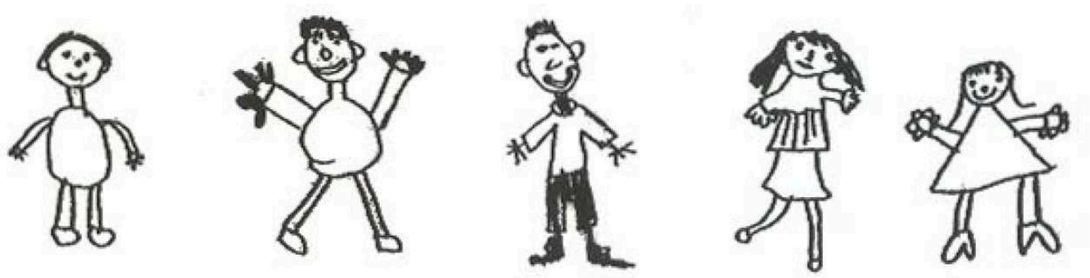


Abbildung 3.8 Zeichnungen der Wenigseher (Quelle: Winterstein (2006))

Mit im Schnitt über zehn Punkten erreichen die Kinder die besten Ergebnisse, die am wenigsten Zeit, maximal 60 Minuten, vor dem Bildschirm verbringen. Abbildung 3.8 präsentiert typische Menschzeichnungen dieser Vorschulkinder. Die nachvollziehbaren Proportionen, Arme und Hände, Beine und Füße, Köpfe und Augen sind eindeutig zu erkennen.

Wie die nachfolgende Abbildung 3.9 bereits vermuten lässt, handelt es sich hierbei um die Zeichnungen derjenigen Vorschulkinder, die sehr viel Zeit vor dem Fernsehgerät verbringen. Ab einer täglichen Fernsehdosis von drei Stunden bestehen die „Figuren" oft nur noch aus einem Kopf, einem Rumpf und einigen Strichen.

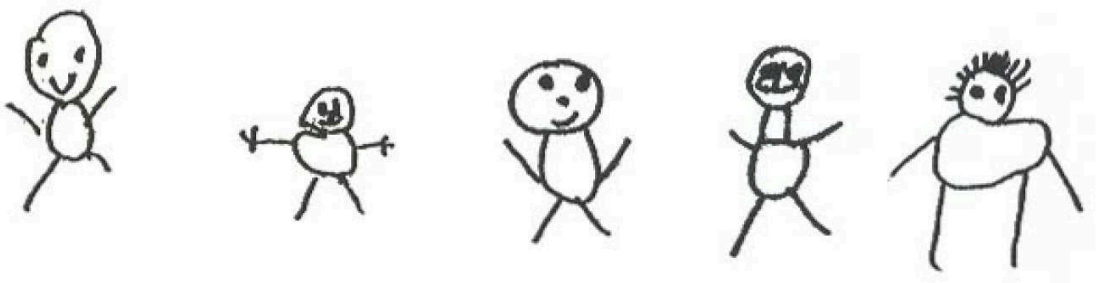


Abbildung 3.9 Zeichnung der Vielseher (Quelle: Winterstein (2006))

Während die optischen Ergebnisse bereits auf imposante Weise verdeutlichen, inwieweit die bildhafte Wahrnehmung durch den Fernsehkonsum beeinflusst bzw. beeinträchtigt werden kann, unterstreicht die streng monoton fallende Kurve in Abbildung 3.10 offensichtlich, dass die Punkte im *Mensch-Zeichen-Test* mit zunehmender Fernsehdauer gravierend fallen, so dass WINTERSTEIN feststellt: *„Es fand sich eine hoch signifikante Dosis-Wirkungs-Beziehung zwischen täglicher Fernsehdauer und den Ergebnissen im MZT."* (ebd., S. 4)

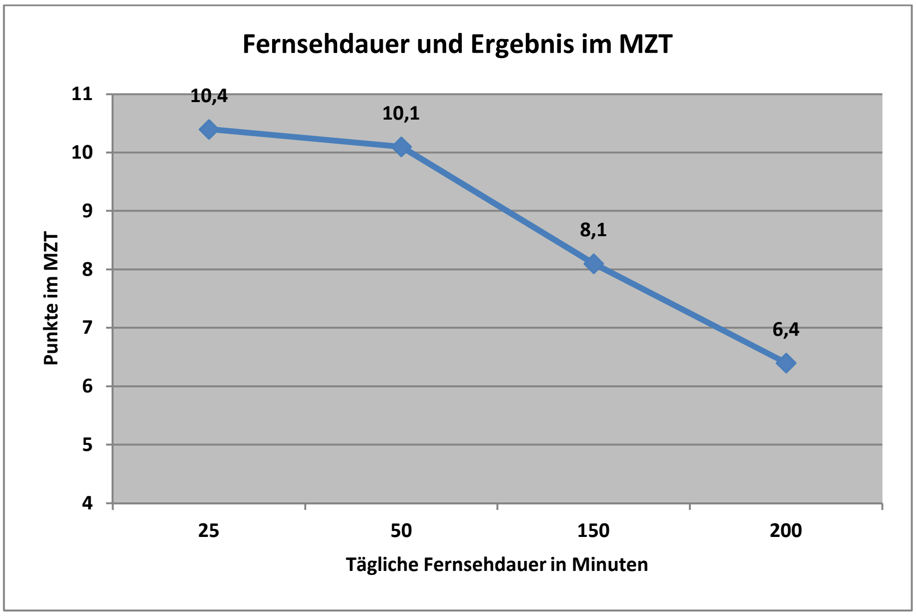


Abbildung 3.10 Durchschnittliche Ergebnisse im „Mensch-Zeichentest" gruppiert nach der Dauer des täglichen Fernsehkonsums (Quelle: Winterstein (2006))

Zusätzlich zu den bis hierhin vorgestellten Ergebnissen, ermöglicht es der *Mensch-Zeichen-Test* auch, Rückschlüsse bezüglich einer korrekten Mengenerfassung der Vorschulkinder zu ziehen. So setzt eine richtige Darstellung der Hände mit jeweils fünf Fingern logischerweise zumindest ein gewisses Mengenverständnis voraus. Tatsächlich ist es denjenigen Kindern, die maximal 90 Minuten täglich fernsehen in 71,6 % aller Fälle gelungen, Hände mit der richtigen Fingerzahl zu zeichnen, während gerade mal 25,5 % der Vielseher (mindestens 90 Minuten täglich) dazu in der Lage waren. Folgerichtig lässt sich auf Basis des *Mensch-Zeichen-Tests* nicht nur ein negativer Zusammenhang zwischen der bildlichen Wahrnehmung und dem Fernsehkonsum von Kindern feststellen, sondern auch ein negativer Zusammenhang zwischen der Mengenerfassung und dem Fernsehkonsum. Dies ist sicherlich ein Aspekt, der für den Mathematikunterricht in der Grundschule von elementarer Bedeutung ist (vgl. ebd., S. 4).

Nicht verwunderlich also, dass Peter Winterstein (2006, S. 6) zum Ende seiner Ausführungen dazu aufruft, den Medienkonsum bei Kindern zeitlich einzuschränken.

3.4.3 Das Sesamstraßenexperiment

Angesichts der zahlreichen Studien, die einen negativen Einfluss des Fernsehens auf die kindliche und jugendliche Entwicklung tatsächlich zu bestätigen scheinen, überrascht es durchaus, dass sich auch einige empirische Untersuchungen finden lassen, die auf positive Effekte eines gemäßigten und geregelten Fernsehkonsums hindeuten. Selbst die Wissenschaftler sind sich folgerichtig bei der Bewertung des Mediums uneins. Da sich die meisten negativen Studien vor allem auf den Umstand beziehen, dass besonders viel fernsehen schadet, liegt es nahe, dass sich positive Studien, exemplarisch wird im Folgenden auf verschiedene empirische Untersuchungen zur Wirkungskraft der *Sesamstraße* eingegangen, eher mit den Fragen *was wird geschaut* bzw. *was sollte geschaut werden* auseinandersetzen. Dieser Ansatz wird u. a. auch von Ingrid Paus-Hasebrink (2010), Kommunikationswissenschaftlerin der Universität Salzburg, vertreten. *„Wenn man den Fokus nur auf den Fernsehkonsum legt, macht man einen Fehler"*, kritisiert sie in der Süddeutschen Zeitung.

3.4.3.1 Das Sesamstraßenexperiment in Bangladesch

Um zu erforschen, ob das Fernsehen positive Auswirkungen auf die Bildung von Kindern haben kann, startete das Institut Associates for Community and Population Research 2003 eine interessante Untersuchung. Durchgeführt wurde diese allerdings nicht in Westeuropa oder den USA, sondern in Bangladesch. Die Wissenschaftler sahen hier die einmalige Chance, aussagekräftige Ergebnisse über die Wirkung des Fernsehens sammeln zu können, da in Bangladesch einerseits das Bildungsniveau äußerst nachhilfebedürftig ist, die Analphabetenquote liegt beispielsweise bei etwa 60 %, aber andererseits lediglich 45 % aller Familien ein Fernsehgerät besitzen, so dass der Fernsehkonsum und die Programmvielfalt deutlich geringer als etwa in Deutschland sind. Zum Vergleich sei angeführt, dass 99 % der deutschen Haushalte mit einem Fernsehgerät ausgestattet sind und nach Schätzung des Bundesverbands für Alphabetisierung und Grundbildung *nur* 4 % der Deutschen als Analphabeten angesehen werden (vgl. Bundesverband für Alphabetisierung und Grundbildung o.J.).

Für das Forschungsprojekt sollte die Wirkung der *Sesamstraße* untersucht werden, einer Sendung, die sich auch in Deutschland selbst das Ziel setzt, zur Verbesserung der Bildung beizutragen und zu beweisen, dass Lernen Spaß machen kann. Aus diesem Grund wurden

2003/2004 in Zusammenarbeit mit bangladescher Bildungsexperten und dem Kreativteam von Sesame Workshop Bildungsziele für das Projekt herausgearbeitet, die als Grundlage für das pädagogische Konzept der Serie dienen und dem Kulturkreis angepasst werden sollten. Die Regierung befürwortete das Projekt als eine preiswerte bildungspolitische Maßnahme, die relativ viele Kinder im Land erreichen könne (vgl. LEE 2008, S. 50 f).

Im April 2005 wurde die Sendung erstmals in Bangladesch unter dem Titel *Sisimpur* ausgestrahlt. Das bekannte Motto der Kindersendung lautet in Deutschland seit über 40 Jahren „Der, Die, Das! - Wer, Wie, Was? - Wieso, Weshalb, Warum? - Wer nicht fragt, bleibt dumm!". Die halbstündigen Folgen setzen sich aus lehrreichen Puppendialogen („Ein Kreis ist rund", „Dieser Buchstabe ist ein A"), Trickfilmen, Kinderliedern und Realfilmbeiträgen über einfache Situationen aus dem Kinderalltag zusammen und folgen damit dem „Magazin"-Format. Dadurch wird vor allem das Ziel verfolgt, die Lese-, Schreib- und Rechenfähigkeiten von Vorschulkindern zu verbessern. Zusätzlich liegt ein Fokus von *Sisimpur* aber auch auf der Vermittlung von sozio-kulturellen Kompetenzen, so dass sich zahlreiche Beiträge mit Hygiene, Respekt, Toleranz, Familie, sozialen Beziehungen und der bangladescher Kultur auseinandersetzen (vgl. LEE 2009, S. 5 f).

Um zu überprüfen, ob die bildungspolitischen Ziele tatsächlich zu erreichen sind, brachten speziell ausgestattete Rikschas Fernseher selbst in die abgelegenen, fernsehfreien und bildungsfernen Dörfer und ermöglichten landesweit den Kindern die Serie regelmäßig zu schauen. Viele Kinder sahen dabei zum ersten Mal fern. Insgesamt wurde die pädagogisch geprägte Sendung so gut aufgenommen, dass innerhalb kürzester Zeit sowohl der Bekanntheitsgrad von *Sisimpur* anstieg, als auch die Beliebtheit der Sendung. Mittlerweile geben beispielsweise 55 % der Kinder an, dass *Sisimpur* ihre Lieblingssendung im Fernsehen sei. Sogar 75 % der Kinder bestätigen zudem, dass *Sisimpur* ihre am häufigsten gesehene Sendung ist (vgl. ebd., S. 16).

Zur Analyse des möglichen Einflusses der Serie auf die Bildung der bangladescher Kinder wurde in einer ersten Welle 2006 und in einer zweiten Welle 2007 jeweils beobachtet, ob die Kinder, welche die Sendung regelmäßig schauten über höhere Kompetenzen in den drei Bereichen *Lese- und Schreibfähigkeiten*, *Mathematik* und *Sozio-Kulturelles* verfügen als Kinder, die die Sendung nicht oder nur selten gesehen haben. Die 7.000 vier- bis siebenjährigen an der Studie teilnehmenden Kinder wurden dazu entsprechend ihres *Sisimpur*-Konsums in drei Gruppen eingeteilt und nach verschiedenen grundlegenden Fähigkeiten beurteilt. Im Bereich der *Lese- und Schreibfähigkeiten* wurden beispielsweise die Größe des Wortschatzes und das Erkennen von Buchstaben getestet, während die *mathematischen* Tests sich auf das Zählen und das Erkennen von Zahlen spezialisierten. Die *sozio-kulturellen* Untersuchungen zielten dagegen auf geografische, kulturelle und soziale Kenntnisse ab. Folgende Ergebnisse konnten von den Forschern in den drei Testbereichen festgestellt werden (vgl. ebd., S. 22 ff):

(i) Lese- und Schreibfähigkeiten

Der Zusammenhang zwischen den Testergebnissen im Bereich der Lese- und Rechtschreibfähigkeiten und dem Konsum von *Sisimpur* wird in den Abbildungen 3.11 (Welle 1) und 3.12 (Welle 2) dargestellt. Auf der x-Achse des Säulendiagramms wird dazu das Alter der vier- bis siebenjährigen Probanden angegeben, während die Werte auf der y-Achse die im Test erzielten Durchschnittspunktzahlen zwischen 0 und 100, jeweils getrennt für Wenig- Mittel- und Vielseher von *Sisimpur*, repräsentieren.

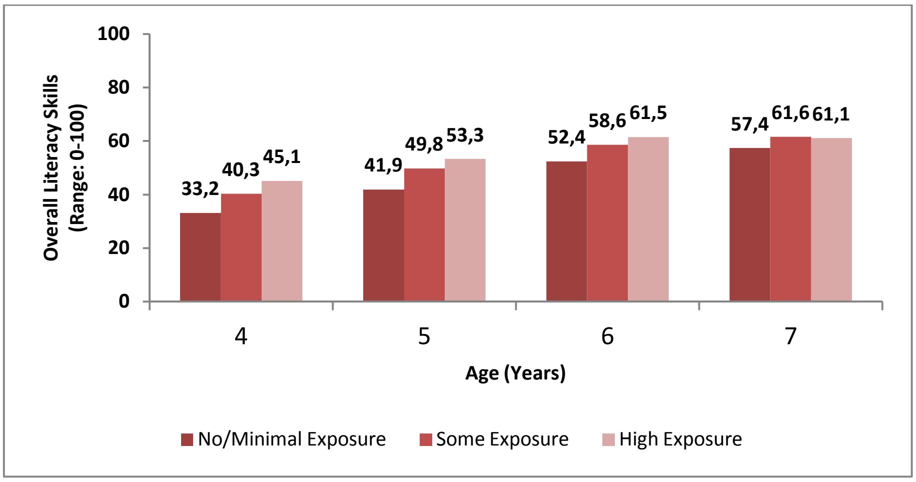


Abbildung 3.11 Lese- und Rechtschreibfähigkeiten Welle 1 (Quelle: Lee (2008))

Wie die erste Abbildung unterstreicht, fanden die Forscher heraus, dass das Anschauen von *Sisimpur* mit höheren Lese- und Schreibfähigkeiten verbunden ist. Es wird deutlich, dass bereits bei der ersten Auswertung 2006 die Kinder, die die Sendung am häufigsten geschaut haben in ihrer Entwicklung teilweise ein ganzes Jahr weiter sind, als die Kinder, die die Sendung seltener oder gar nicht gesehen haben. Allerdings lässt sich zurecht einwenden, dass der Effekt im Laufe der Zeit deutlich nachlässt. Bei den siebenjährigen Kindern zeigen beispielsweise die Kinder, die *Sisimpur* nur hin und wieder sehen, die besten Leistungen.

Betrachtet man allerdings nun die Ergebnisse der zweiten Auswertung 2007, lässt sich auch dieser Kritikpunkt entkräften. Scheinbar hat die kontinuierliche Überarbeitung und Verbesserung der Sendung dazu geführt, dass auch mit steigendem Alter diejenigen Kinder, die die Sendung oft schauen, mit deutlichem Abstand die besten Fähigkeiten im Bereich des Lesens und Schreibens besitzen (vgl. ebd., S. 23 ff).

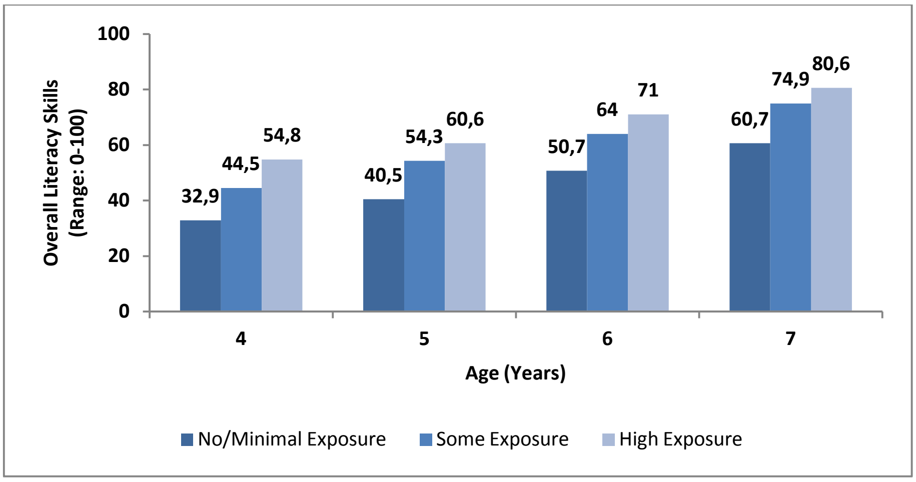


Abbildung 3.12 Lese- und Rechtschreibfähigkeiten Welle 2 (Quelle: Lee (2008))

(ii) Mathematische Fähigkeiten

Hinsichtlich der mathematischen Fähigkeiten konnte das Forschungsteam nachweisen, dass die Kinder in Bangladesch vor der Einführung von *Sisimpur* im Schnitt bis 18,2 zählen und 3,0 Zahlen aus einer Auswahl von 6 Zahlen identifizieren konnten. Diese Durchschnittswerte konnten auch mit Hilfe der neuen Sendung gesteigert werden. Mittlerweile zählen die Kinder im Schnitt aller vier- bis siebenjährigen bis 25,3 und erkennen 3,6 aus 6 Zahlen (vgl. ebd., S. 24 ff).

Betrachtet man nun die beiden Säulendiagramme, die analog zu den Lese- und Rechtschreibdiagrammen die Zusammenhänge zwischen den mathematischen Testergebnissen und dem Konsum von *Sisimpur* darstellen, lassen sich ähnlich signifikant positive Forschungsergebnisse identifizieren:

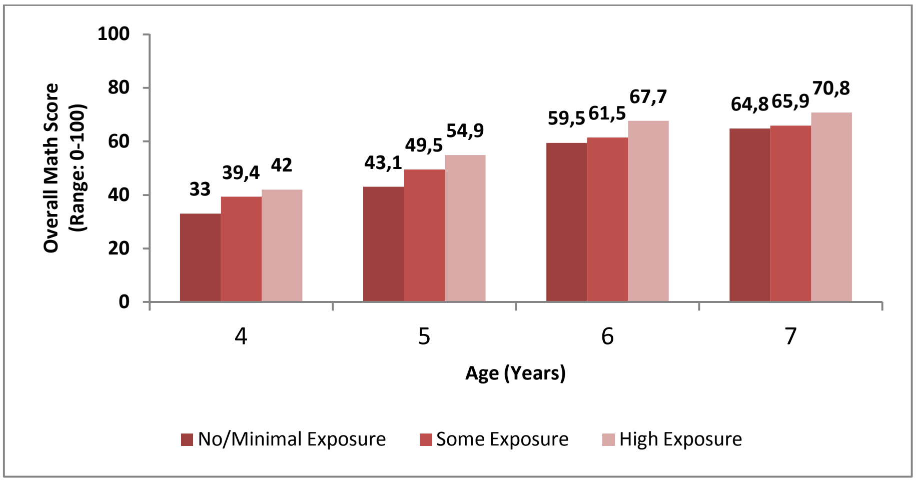


Abbildung 3.13 Mathematische Fähigkeiten Welle 1 (Quelle: Lee (2008))

Gerade diejenigen Probanden, welche die Sendung besonders häufig sehen, weisen in jeder Altersgruppe die mit Abstand besten mathematischen Fähigkeiten auf, wie man Abbildung 3.13 sofort entnehmen kann. Der Leistungsunterschied zwischen den Mittel- und Wenigsehern nimmt in der ersten Testwelle mit zunehmendem Alter allerdings ab, wobei diese Beobachtung nach Auswertung der zweiten Welle (Abbildung 3.14) in Frage gestellt werden darf. Die Verbesserungen der Probanden nehmen in dieser zweiten Welle generell noch einmal deutlich, im Schnitt um 10%, zu. Die Abstinenzler sind also auch in ihrer mathematischen Entwicklung teilweise ein ganzes Jahr im Rückstand (vgl. ebd., S. 24 ff).

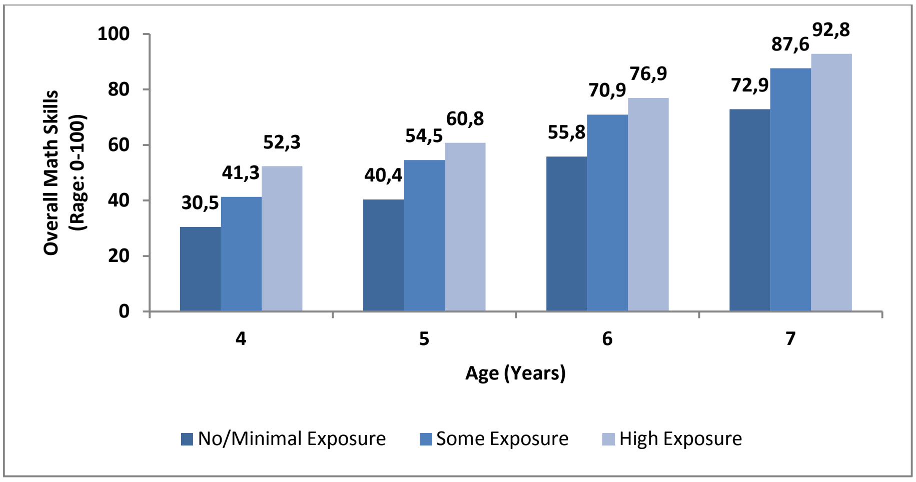


Abbildung 3.14 Mathematische Fähigkeiten Welle 2 (Quelle: Lee (2008))

(iii) Sozio-Kulturelle Fähigkeiten

Die Ergebnisse im sozio-kulturellen Bereich unterstreichen abermals die schon vorgestellten Auswertungen. Daher soll an dieser Stelle auf eine tiefer gehende Analyse verzichtet werden. Angemerkt werden sollte allerdings dennoch, dass viele Fernsehkritiker gerade in diesem Bereich dem Fernsehen vorwerfen, zu versagen und den Kindern in ihrer Entwicklung zu schaden. Auch diese Kritik wird durch das *Sesamstraßenprojekt* zumindest in Frage gestellt (vgl. ebd., S. 25 ff).

3.4.3.2 Weitere Studien zur Sesamstraße

Das *Sesamstraßenexperiment* in Bangladesch ist die aktuellste und modernste Untersuchung, welche positive Bildungseffekte von pädagogisch geprägten Fernsehsendungen herausarbeitet. Doch bereits 1969, kurz vor der offiziellen Einführung der *Sesamstraße* in Amerika, wurden anhand von Testsendungen in Kindertagesstätten mehrere Untersuchungen bezüglich eines möglichen Lerneffekts durch die *Sesamstraße* durchgeführt. Eine Episode beinhaltete beispielsweise einen 60-Sekunden-Beitrag zu dem Buchstaben „J". Nun wurden die Kinder in verschiedene Gruppen eingeteilt, die unterschiedlich oft diesen sogenannten „Jot-Spot" sahen. Es zeigte sich, dass diejenigen Kinder, die diesen Sport sechsmal gesehen haben, im Anschluss an die Sendung am ehesten in der Lage waren, das „J" wiederzugeben und aus einer Menge an Buchstaben herauszufiltern. Die Entwickler der *Sesamstraße* schlussfolgerten, dass regelmäßige Wiederholungen von solchen kurzen Beiträgen wertvoll sind, damit die Kinder die darin enthaltenen Informationen besser abspeichern können. Folgerichtig werden noch heute Buchstaben und Zahlen auch in den Zeichentrickfilmbeiträgen immer wieder verwendet, um das kinetische Lernen der Kinder zu fördern (vgl. FEINSTEIN 1971, S. 50).

Zu Beginn des Jahres 1970 wurden Resultate weiterer Voruntersuchungen in Amerika veröffentlicht. Eine dieser Studien befasste sich mit 130 sozial benachteiligten Kindern im Vorschulalter, die in zwei Gruppen eingeteilt wurden. Probanden der Gruppe I sahen die *Sesamstraße* regelmäßig, Probanden der Gruppe II sahen die Sendung nie. Es konnte gezeigt werden, dass diejenigen Kinder, die mindestens sechs Wochen lang die *Sesamstraße* verfolgten, mehr als doppelt so viel lernten als die Kinder, welche die Sendung nicht sahen. Gerade für Kinder aus sozial schwächeren Familien schien der Lerneffekt besonders groß zu sein (vgl. ebd., S. 88). Dass die Schichtzugehörigkeit den Lernerfolg nicht einzuschränken vermag und die *Sesamstraße* allen Kindern offen steht, bestätigt auch eine weitere Studie, die zusammenfassend festhält, dass der Bildungseffekt unabhängig davon ist, ob es sich bei dem zusehenden Kind um ein *„unter-privilegiertes Innenstadtkind, privilegiertes Vorstadtkind, isoliertes Landkind, Junge oder Mädchen und Kinder, deren erste Sprache nicht Englisch ist"* (ebd., S. 145) handelt.

Neben Studien zum Lernerfolg bei Buchstaben wurden vom Educational Testing Sevice zusätzlich weitere Tests zur Überprüfung des erworbenen Wissens zu Zahlen, Formen, Klassifikationen und Körperteilen anhand von 943 Kindern durchgeführt (vgl. ebd., S. 144). Auch hier konnte belegt werden, dass der Lerneffekt bei Kindern in Abhängigkeit von der Anzahl und der Regelmäßigkeit der gesehenen Sendungen steigt. Darüber hinaus ist von Bedeutung, welche inhaltlichen Einzelaspekte innerhalb des Programms fokussiert werden, denn die Aspekte, denen die größte Aufmerksamkeit geschenkt wurden, lernten die Kinder am besten (vgl. SCHLEICHER 1972, S. 28).

Anhand von 570 Probanden wurden zudem die Langzeiteffekte der *Sesamstraße* auf die weiteren Schulleistungen untersucht. Zunächst wurde der Fernsehkonsum im Alter von fünf Jahren dokumentiert. Der Fernsehnutzungsanteil im Kindesalter, der der *Sesamstraße* zugeschrieben wurde, lag dabei zwischen 1,7 und 2,7 Stunden in der Woche. Als dieser später in Zusammenhang mit den Noten in der Highschool gebracht wurde, zeigte sich, dass die Jungen, die im Kindesalter häufig (fünfmal pro Woche) die *Sesamstraße* sahen, im Notendurchschnitt besser abschnitten als die Jungen, die diese Sendung selten oder gar nicht sahen. Bei den Mädchen blieb dieser Effekt allerdings weitestgehend aus (vgl. ENNEMOSER 2003a, S. 70).

3.4.3.3 Fazit zum Sesamstraßenexperiment

Die durchweg positiven Ergebnisse sowohl der modernen als auch der „historischen" Studien, unabhängig vom Alter, dem Geschlecht und der sozialen Herkunft der teilnehmenden Kinder, können sicherlich begeistern, vor allem wenn man bedenkt, dass die Wirkung nachhaltig über einen Zeitraum von zwei Jahren in Bangladesch nachgewiesen werden konnte. Allerdings ist es dennoch diskutabel, inwieweit die aktuellen Ergebnisse aus Bangladesch auf Deutschland und das westliche Fernsehen im Allgemeinen übertragbar sind. Schließlich handelt es sich um eine *einzige* Sendung in einem *einzigen* Land, welches sich kulturell, politisch und wirtschaftlich stark von Deutschland unterscheidet. Zudem sollte beachtet werden, dass sich in Westeuropa und den USA die Mediennutzgewohnheiten seit der Einführung der *Sesamstraße* 1969 stark verändert haben. Die Eltern finden sich in der verantwortlichen Rolle wieder, diesem Trend entgegenzuwirken und ihren Kindern derartige Sendungen schmackhaft zu machen; zeigt das Ergebnis aus Bangladesch doch eindeutig, dass Kinder in gut gemachten und pädagogisch ausgerichteten Fernsehsendungen nicht nur Fakten aufnehmen, sondern auch etwas lernen können.

3.5 Forschungsfragen und Hypothesen

Nachdem im zweiten Kapitel eindrucksvoll belegt werden konnte, dass das Fernsehen sowohl im Kindes- und Jugendalter als auch innerhalb des täglichen Familienlebens nach wie vor das dominierende Medium ist, stand in dem dritten Kapitel der daraus resultierende Einfluss auf die Bildung bzw. den Schulerfolg von Schülerinnen und Schülern im Mittelpunkt.

Das visuelle Medium Fernsehen *kann dumm* machen, muss es aber nicht. Nach der Gegenüberstellung der jeweiligen Meinungen und Untersuchungen scheint dies die triviale Erkenntnis über den Einfluss der Bildschirmmedien auf die Bildung zu sein.

Gegner als auch Befürworter sind sich insoweit einig, dass der Umgang mit dem Medium entscheidend ist. Die kritischen Studien prangern vor allem an, dass Kinder zu früh eigene Geräte besitzen, diese unreflektiert nutzen und zu viel konsumieren. In keiner Studie ist zu lesen, dass die *Sesamstraße* oder *Die Sendung mit der Maus* jemals einem Kind geschadet haben. Doch die Fähigkeit, mit den Medien verantwortungsbewusst umzugehen, müssen die Kinder erlernen. Die Schule sowie vor allem das Elternhaus spielen hier die zentrale Rolle und haben die Aufgabe, mit den Kindern Fernsehsendungen auszuwählen, bei denen etwas gelernt werden kann oder ein wöchentliches Arrangement über die Fernsehdauer zu treffen. Zudem müssen die Eltern ihre Kinder aber auch dazu ermutigen, sich aktiv in der übrigen Freizeit zu betätigen. Das Zitat von Christian Pfeiffer (2005) bringt die Sicht der wissenschaftlichen Kritiker auf den Punkt:

> *„Also, wir sind keine Bilderstürmer, die behaupten: 'Fernsehen und Computer sind des Teufels!' Das sind sie nicht. Es gibt wunderbare Filme, die zu sehen, einen persönlich weiterbringen können, die Diskussionen auslösen, die einen unterhalten, die den Spaß im Leben erhöhen. Daher suchen wir eine realistische Perspektive zu der Frage: Was ist ein vernünftiger, sich gut ins Leben integrierender Konsum von Computerspielen und Fernsehen?"* (zitiert nach KRELL 2005)

Die Befürworter der Bildschirmmedien sind sich einig, dass pädagogisch wertvolles Fernsehen der kindlichen und jugendlichen Entwicklung förderlich sein kann. In keiner Studie wird aber erwähnt, dass es hilfreich ist, dauernd und vor allem in sehr frühem Alter vor dem Bildschirm zu sitzen. Damit das Fernsehen sein förderliches Potenzial tatsächlich entfalten kann, sollte es nur gemäßigt genutzt werden. Während des Konsums sollte die Aufmerksamkeit stabil bleiben. Lediglich nebenbei konsumierte Sendungen können keinen positiven Einfluss auf die kognitive Entwicklung der Kinder haben.

Bei der Analyse der Studien ist dem Verfasser zweierlei aufgefallen: Einerseits berufen sich die Befürworter der Bildschirmmedien vor allem auf den wertvollen Inhalt von Fernsehsendungen, während die Kritiker sich eher mit der Fernsehdauer auseinandersetzten. Des Weiteren wird in den meisten Studien das Fernsehverhalten vor allem in Bezug zur Entwicklung der Lese- und Schreibkompetenzen von Kindern und Jugendlichen gesetzt. Obwohl die Mathematik die Menschen ebenfalls von Tag zu Tag durch das Leben begleitet und die Teilnahme am gesellschaftlichen Leben ermöglicht, spielt das Fachgebiet in den vorliegenden Untersuchungen eher eine untergeordnete Rolle.

Aus diesem Grund konzentriert sich im Folgenden der eigenständig konzipierte Feldversuch darauf, einen möglichen Zusammenhang zwischen Fernsehkonsum einerseits und Rechenfähigkeiten andererseits zu ermitteln. Dazu werden Siebtklässler an einem Gymnasium, zwei Gesamtschulen und einer Hauptschule fragebogenbasiert zu ihrer Fernsehdauer, ihren Lieblingssendungen und ihrem Freizeitverhalten befragt. Die Ergebnisse werden mit den erhobenen Mathematikleistungen in Bezug gesetzt.

Die zentrale Forschungsfrage lautet dementsprechend: Bestehen Zusammenhänge zwischen Fernsehverhalten einerseits und von Lehrerinnen und Lehrern beurteilter Mathematikleistung andererseits bei Schülerinnen und Schülern der Sekundarstufe I?

Auf Basis, der aus der Literatur bekannten Forschungsergebnisse stellt der Verfasser die Hypothese auf, dass ein überhöhter zeitlicher Fernsehkonsum und ein passives und medial geprägtes Freizeitverhalten mit unterdurchschnittlichen Leistungen in Mathematik verbunden sind, während das Fernsehprogramm in einer positiven Beziehung zu den erbrachten Leistungen stehen kann.

Zu beantwortende Forschungsfragen betreffen aber auch die Methodik der Datenanalyse, da im Fokus der nachfolgenden Untersuchung zudem das zusätzliche Ziel steht, eine wirkliche Einheit von empirischer Methodik einerseits und didaktischer Interpretation andererseits zu schaffen. Es soll also zwingend vermieden werden, auf statistische Standardmethoden zurückzugreifen, und sich das zu interpretierende Ergebnis von dem gewählten datenanalytischen Verfahren diktieren zu lassen. Die entsprechende Forschungsfrage lautet daher: Mit welchen bestehenden oder neu zu entwickelnden datenanalytischen Verfahren können die zu erhebenden Daten adäquat ausgewertet werden?

Da im Verlauf des Forschungsprojekts Daten auf Basis von *Einschätzungsskalen* gemessen werden, muss davon ausgegangen werden, dass die Daten nicht notwendig intervallskaliert sind. Dementsprechend stellt der Verfasser die Hypothese auf, dass die vorhandenen multivariaten Standardverfahren, die gewöhnlicherweise bei intervallskalierten Daten eingesetzt werden, modifiziert und weiterentwickelt werden müssen. Zu diesem Zweck führt zunächst ein mathematischer Teil sowohl in die grundsätzliche als auch in die besondere Problematik empirischer Untersuchungen in der mathematikdidaktischen Forschung ein.

Methodischer Teil

4. Mathematische Grundlagen

4.1 Forschungsansätze zur Behandlung nicht notwendig intervallskalierter Daten in den empirischen Wissenschaften

Bevor mit der Datenerhebung und der Auswertung einer empirischen Studie begonnen werden kann, gilt es vorab zu dem Niveau der erhobenen Daten angemessene datenanalytische Verfahren und Methoden auszuwählen oder falls nötig, zu entwickeln. Werden standardisierte Verfahren, die sich in der Vergangenheit bei ähnlichen Fragestellungen mutmaßlich bewährt haben, verwendet, gehen Anwender das Risiko ein, dass nicht länger der tatsächlich befragte Proband für die Struktur der Daten verantwortlich ist, sondern die Methode. Der Erfolg und der Gehalt der inhaltlichen Interpretation der Ergebnisse eines Forschungsprojekts hängen dementsprechend in großem Maße davon ab, ob sich die verwendeten Auswertungsverfahren für das Messniveau (Skalenniveau) der Daten eignen.

Die vier wichtigsten Skalenarten sind laut Fahrmeir und Hamerle (1996, S. 8) die Nominalskala, die Ordinalskala, die Intervallskala und die Verhältnisskala. In dieser Arbeit werden vor allem Daten erhoben und ausgewertet, die nicht notwendig intervallskaliert sind. Dabei handelt es sich im Folgenden um Daten, deren Messniveau (Skalenniveau) durch eine Transformationsgruppe beschrieben wird, die nicht ausschließlich aus Transformationen der Form $x \mapsto ax + b$ besteht. Im Hinblick auf das in der Praxis empirischer Forschung in der Regel sekundäre Forschungsziel, der Entwicklung von für die Analyse nicht notwendig intervallskalierter Daten geeigneter Methoden, lassen sich die folgenden drei Forschungsansätze, die sowohl den aktuellen Stand der Forschung widerspiegeln, als auch die entsprechende Methodendiskussion beherrschen, unterscheiden.

4.1.1 Forschungsansatz nach Jardine, Sibson und Janowitz

Der erste Forschungsansatz geht auf Jardine und Sibson (1971) zurück und beschränkt sich (im Wesentlichen) auf Methoden der *Numerischen Taxonomie* mit dem Ziel Klassifikationsprozesse mathematisch *„zu 'objektivieren' und die Güte der Klassifikation überprüfbar zu* *machen"* (Hohenberger 1982, S. 16). Dieser Forschungsansatz wurde noch in den siebziger Jahren des vorigen Jahrhunderts von Janowitz und seiner Schule aufgegriffen und zu einer vollständigen, auf rein ordnungstheoretischen Methoden der Mathematik basierenden, Theorie hierarchischer Verfahren der Clusteranalyse weiterentwickelt. Eine Zusammenfassung der wichtigsten Ergebnisse dieses Forschungsansatzes liefert das entsprechende Buch von Janowitz (2010).

4.1.2 Die formale Begriffsanalyse

Der zweite und vielleicht bemerkenswerteste Forschungsansatz hat seit den späten siebziger und den frühen achtziger Jahren des vergangenen Jahrhunderts seinen Ursprung in der von Wille begründeten und wenig später von Ganter mitbegründeten Schule der formalen Begriffsanalyse. Ihr ebenso einfacher wie eleganter Ansatz beruht auf den Konzepten des formalen Kontexts und des formalen Begriffs sowie der fundamentalen Feststellung, dass die formalen Begriffe eines Kontexts in natürlicher Weise einen vollständigen Verband bilden. Methodisch lässt sich dieser Ansatz also der Verbandstheorie zuordnen, was eine enge methodische, nicht unbedingt inhaltliche, Nähe zu dem von Janowitz begründeten Forschungsansatz bedeutet. Die Breite der Anwendungsmöglichkeiten der formalen Begriffsanalyse ist kaum zu überschätzen. Sie betrifft nicht nur die Analyse gewonnener Daten, sondern auch (weite) Teile der reinen Mathematik. Das erste Buch zur formalen Begriffsanalyse geht auf Wille und Ganter (1996) zurück.

4.1.3 Standardisierte Methoden in den Sozialwissenschaften

Obwohl die bisher aufgeführten Forschungsansätze für die Analyse nicht notwendig intervallskalierter Daten gut geeignet sind, stoßen die beiden Ansätze in den empirischen Sozialwissenschaften nicht auf die ihrer Bedeutung entsprechende Resonanz. Vielmehr werden in den empirischen Sozialwissenschaften selbst bei *Einschätzungsskalen* mögliche Abweichungen von einer Intervallskala häufig einfach ignoriert, in dem behauptet wird, dass auch diese Skalen wie intervallskaliert behandelt werden dürfen (man vergleiche hierzu das im Internet zur Verfügung stehende Skriptum von Payrhuber, die am Institut für Publizistik und Kommunikationswissenschaft der Universität Wien lehrt und forscht, mit dem Titel „Auswertung und Darstellung wissenschaftlicher Daten 2"). In diesem Zusammenhang schreiben Bortz und Döring (2006, S. 70) beispielsweise, dass in der Praxis der empirischen Forschung sogar häufig bewusst auf eine genaue Analyse der Skalenaxiomatik verzichtet wird, um sämtliche statistische Auswertungsmethoden für Intervallskalen anwenden zu können. In diesen Fällen spricht man von *Per-Fiat-Messungen*, was bedeutet, dass die pragmatischen Anwender darauf vertrauen, dass mögliche Verletzungen der Intervallskaleneigenschaften im Sinne des Einsatzes der standardisierten Methoden in Kauf genommen werden können. Es wird dementsprechend vorausgesetzt, dass Probanden beispielsweise die Antwortalternativen bei *Ein-*

schätzungsskalen als äquidistant auffassen (vgl. LANDAUER 1997). Man vergleiche hierzu auch die Bücher von FISCHER (1991) und BACKHAUS et al. (1996).

Im Gegensatz dazu begegnen empirisch forschende Sozialwissenschaftler, die sich mit der Skalenaxiomatik auseinandersetzen, einem Messniveau (Skalenniveau), welches *schwächer* ist als das Niveau einer Intervallskala, durch Modifikation und Erweiterung klassischer Methoden der Datenanalyse, die an dichotome Variable, multiple kategoriale Variable, ordinale Variable oder Zählvariable angepasst ist. Dies meint, dass durch spezielles Eingehen auf die individuelle Datenstruktur der Versuch unternommen wird, die bekannten Analyseverfahren unabhängig vom Skalenniveau verwenden zu können. Folglich wird in den empirischen Sozialwissenschaften bei der Analyse nicht notwenig intervallskalierter Daten eher konservativ vorgegangen. Detaillierte Informationen können bei AGRESTI (1996) und BENNINGHAUS (2005) nachgelesen werden.

4.1.4 Methodischer Ansatz dieser Arbeit

Vor dem Hintergrund der aufgeführten Forschungsansätze kann das primäre Forschungsziel in dieser Dissertation nicht darin bestehen, zu den kaum überblickbaren Verfahren zur Analyse nicht notwendig intervallskalierter Daten weitere noch besser geeignete und selbstverständlich erfolgreichere Verfahren hinzuzufügen. Vielmehr besteht das Ziel dieser Arbeit darin, bekannte Standardverfahren so zu modifizieren und weiterzuentwickeln, dass sie datenadäquat angewendet werden können. Aus diesem Grund steht das vierte Kapitel „Mathematische Grundlagen" ganz im Zeichen einer ausführlichen Methodenreflexion. Zunächst wird ein Vorschlag zur Behandlung der Kompatibilitäts- und Adäquatheitsproblematik faktorenanalytischer Verfahren präsentiert und im Anschluss diskutiert, wie die Probanden entsprechend des konkreten Antwortverhaltens auf die verschiedenen durch die Faktorenanalyse identifizierten Faktoren aufgeteilt werden können. Durch die Erweiterung des Hauptsatzes der formalen Begriffsanalyse auf beliebige Ordnungen soll auf die Methodenunabhängigkeit von Forschungsergebnissen hingewiesen werden. Abschließend steht der Umgang mit fehlenden Werten (*missing values*) bei ordinalskalierten Datensätzen im Fokus.

4.2 Ein Vorschlag zur Behandlung der Kompatibilitäts- und Adäquatheitsproblematik faktorenanalytischer Verfahren

Unter den in empirischen Untersuchungen zur Didaktik der Mathematik benutzten multivariaten Methoden der Datenanalyse nehmen Verfahren der Faktorenanalyse, insbesondere die Hauptkomponentenmethode mit anschließender Varimax-Rotation, eine nicht unbedeutende Rolle ein. Dies impliziert schon ein Blick auf die nackten Zahlen der *Mathematics Education Database (Mathematics Didactics Database)*: Unter dem Stichwort *factor analysis* sind zwischen 2000 und 2012 genau 144 Arbeiten erfasst worden, während im gleichen Zeitraum unter den Suchbegriffen *cluster analysis*, *multidimensional scaling* und *Lisrel* bzw. *linear structural equation models* nur 32 (*cluster analysis*), 6 *multidimensional scaling* und 4 (*Lisrel* bzw. *linear structural equation models*) Arbeiten erfasst worden sind. Lediglich unter dem Stichwort *regression analysis* sind seit 2000 mehr Arbeiten (164) zitiert. Nahezu in allen in der *Mathematics Education Database* (*Mathematics Didactics Database*) erwähnten Arbeiten wurden Methoden der Faktorenanalyse im Rahmen empirischer Untersuchungen zur Didaktik der Mathematik benutzt. Nur in den seltensten Fällen setzen sich die Autoren mit den mathematischen Grundlagen von Methoden der Faktorenanalyse auseinander. Zu erwähnen sind hier die Arbeiten von Song und Lee (2002) sowie von Schreiber, King, Stage, Nora und Barlow (2006). Pohlmann (2004) bespricht in seinem Aufsatz 25 Arbeiten über Anwendungen von Verfahren der Faktorenanalyse im Rahmen empirischer Untersuchungen zur Didaktik der Mathematik, die zwischen 1992 und 2002 im *Journal of Educational Research* erschienen sind, und stellt dabei fest, dass in 22 der von ihm untersuchten Artikel Methoden der Faktorenanalyse lediglich als Hilfsmittel der explorativen Datenanalyse eingesetzt wurden und nur in 3 Fällen konfirmatorische Methoden der Faktorenanalyse angewandt worden sind.

Die Zahlen im *MathSciNet* sprechen (auch hier zwischen 2000 und 2012) eine ähnliche Sprache. Aufgelistet sind dort 125 Arbeiten zur Faktorenanalyse, während zur Clusteranalyse (108), zur Multidimensionalen Skalierung (62) und zu Lisrel bzw. Linearen Strukturgleichungen (8) erneut wesentlich weniger Arbeiten zu finden sind. Die Ausnahme bildet auch hier die Regressionsanalyse mit 168 Einträgen.

Unabhängig von diesen Zahlen werden in empirischen Arbeiten zur Didaktik der Mathematik häufig Methoden der Faktorenanalyse (im Wesentlichen die Hauptkomponentenmethode mit anschließender Varimax-Rotation, u. U. ergänzt durch eine Reliabilitätsüberprüfung der Skalen mit Hilfe von Cronbach's α) eingesetzt. Beispielsweise sind die Arbeiten von Grigutsch (1998, 2001) oder Hussmann (2003) zu nennen. In diesen Arbeiten werden Probanden (Schülerinnen und Schüler, Lehrerinnen und Lehrer, usw.) aufgefordert, auf einer *Bewertungsskala* (*Ratingskala*) vorgegebene *Einflussgrößen*, die im Rahmen der vorgenommenen Untersuchung wichtig sind, zu bewerten. Bei einer näheren Beleuchtung der faktorenanalytischen Auswertung und Interpretation in diesen Arbeiten fällt durchaus auf, dass beim Übergang von der Datenmatrix (entspricht der Matrix der beobachteten Zufallsvariablen) zur entsprechenden Korrelationsmatrix ein so hoher Informationsverlust entstehen kann, dass das Ergebnis des Verfahrens der Faktorenanalyse - ohne weitergehende Analysen der Datenstrukturen - im Grunde kaum mehr interpretierbar ist.

Die Erörterung dieser für jede empirische Untersuchung, die faktorenanalytische Methoden benutzt, grundlegende *Kompatibilitätsproblematik zwischen beobachteten Daten einerseits und entsprechender Korrelationsmatrix andererseits* ist das erste Ziel dieses Abschnitts.

Da die bloße Kenntnis und das bloße Bewusstsein dieser Problematik einem Benutzer bei der Anwendung natürlich nur wenig hilft, soll daher in diesem Abschnitt zudem ein konkreter Vorschlag unterbreitet werden, wie sich die erwähnte Kompatibilitätsproblematik in der Praxis empirischer Untersuchungen überwinden lässt. Neben einer Reihe kleinerer Beispiele, deren Berechnungen die mathematische Hintergrundtheorie in vollem Umfang bestätigt haben, soll dieses Prinzip auch in den eigenen größeren empirischen Untersuchungen in den folgenden Kapiteln Anwendung finden. Daher besteht die Sicherheit einen praktikablen Vorschlag dafür anbieten zu können, wie man Verfahren der Faktorenanalyse zur Analyse von mit Hilfe von *Bewertungsskalen* (*Ratingskalen*) ermittelten Daten auswerten und interpretieren kann, der gegenüber den gängigen Verfahren einen erheblichen Fortschritt darstellt. Tatsächlich ist dieser Vorschlag aber nicht nur auf *Bewertungsskalen* (*Ratingskalen*) im engeren Sinne beschränkt, denn natürlich lassen sich sowohl Punktbewertungen im Rahmen von (Mathematik-) Tests als auch kategorisierte Daten als Bewertungen auf einer *Bewertungsskala* (*Ratingskala*) auffassen.

Seit den 1967 auf dem Münchener Symposium vorgetragenen kritischen Arbeiten zur Faktorenanalyse (vgl. Fischer (1967), Kallina (1967), Orlik (1967) und Sixtl (1967)), den häufig diskutierten kritischen Arbeiten von Kalveram (1970a, 1970b, 1970c und 1970d) - man vergleiche hierzu auch Haagen und Oberhofer (1977) - Kempf (1972) und Lukesch und Kleiter (1974) sind mehr als 40 Jahre vergangen, ohne dass die Kritik an der faktorenanalytischen Vorgehensweise kleiner geworden wäre. Ganz im Gegenteil: Die Kritik an den Wissenschaftlerinnen und Wissenschaftlern, die faktorenanalytische Methoden verwenden, wird mit ungeminderter Schärfe weitergeführt (man vgl. hierzu etwa die im Internet unter dem Titel „Internet Publikationen für Allgemeine und Integrative Psychotherapie" (www.sgipt.de) nachlesbaren Arbeiten). Ein Grund dafür ist sicherlich, dass diese Verfahren durch die leichte Verfügbar- und mittlerweile einfache Bedienbarkeit in nahezu allen Gebieten Anwendung finden. Erschreckend ist dabei, dass sich Vorgehensschemata entwickelt haben, die ohne jegliche kritische Reflexion einfach von Arbeit zu Arbeit übernommen werden. Das Ergebnis daraus ist, dass die Software, der es natürlich egal ist, wie die Daten strukturiert sind, auf wie auch immer geartete Daten vorgeschriebene Algorithmen anwendet, die immer zu einem Ergebnis führen. Lohnend also für den geneigten Anwender, denn es kommt schließlich „immer etwas heraus".

Leider kann das Vertrauen in die Richtigkeit eines datenanalytischen Verfahrens den Verlust der Kontrolle über die beobachteten Daten bedeuten. Zusammenhänge werden dann nicht *erforscht, sondern durch die datenanalytische Methode festgelegt. Um das zu verhindern, muss auch die Erörterung und mathematische Reflexion quantitativer Methoden der Datenanalyse im Rahmen empirischer Untersuchungen zu den Forschungsaufgaben der didaktischen Forschung gezählt und konsequent in jedem Einzelfall umgesetzt werden.*

Die in der Literatur zu findende Kritik bezüglich der Faktorenanalyse lässt sich grundsätzlich zu zwei Hauptkritikpunkten zusammenfassen, nämlich zum einen zur Kritik an dem Begriff der *Kommunalität* und zum anderen zur Kritik darüber, dass Rangreduktion (ein Hauptanliegen der Faktorenanalyse) in der Regel lediglich auf Kosten einer nur *unzureichenden Reproduzierbarkeit der Korrelationsmatrix* gelingt. Die Kritik betrifft also *nicht* das Problem des Informationsverlusts beim Übergang der Matrix der beobachteten Daten zur entsprechenden Korrelationsmatrix. Dieses Problem ist der soeben beschriebenen und in der Literatur diskutierten grundsätzlichen Problematik faktorenanalytischer Auswertungen gewissermaßen vorgelagert. Dies bedeutet jedoch nicht, dass das Problem des Informationsverlusts, der Kompatibilität von Datenmatrix einerseits und Korrelationsmatrix andererseits, in der Literatur ignoriert wird. So schreibt z.B. KNOCHE (1990, S. 139):

> *„Allerdings werden diese Vorteile durch einen Informationsverlust erkauft, da man von den konkreten Messergebnissen absieht und unterschiedliche Datenmatrizen zu gleichen empirischen Korrelationsmatrizen führen können".*

Auf der anderen Seite jedoch lassen sich in der Literatur keinerlei Hinweise darauf finden, dass folgende drei, die Problematik des Informationsverlustes beschreibende, Fragen eingehender untersucht worden sind. Dies sind zum Ersten die Frage nach den *möglichen Ursachen* für den Informationsverlust, zum Zweiten die Frage nach dem *möglichen Ausmaß* des Informationsverlusts und zum Dritten die Frage nach der *größtmöglichen Verringerbarkeit* dieses Informationsverlusts. Die Klärung dieser Fragen steht im Mittelpunkt dieses Abschnitts. Im Zusammenhang mit den bekannten und bereits erwähnten empirischen Arbeiten zur Didaktik der Mathematik erscheint ihre Klärung, will man uninterpretierbare faktorenanalytische Auswertungen vermeiden, notwendig.

Obwohl sich die folgenden Überlegungen verallgemeinern lassen, liegt diesem Abschnitt die orthogonale Form des Fundamentaltheorems der Faktorenanalyse zugrunde, nämlich die Gleichung $\operatorname{cov}(X,X) = AA^T + \operatorname{cov}(F,F)$ bzw. $\sum_{XX} = AA^T + E^2$, wobei X und F die Zufallsvektoren der standardisierten beobachteten Variablen bzw. der entsprechenden Fehlervariablen seien und A die zu bestimmende Ladungsmatrix. Zur Umgehung der erwähnten Kritik an der Faktorenanalyse setzen wir im Folgenden voraus, dass $\sum_{XX} - E^2$ als positiv semidefinit vorgegeben ist, und dass die Grundproblematik der Faktorenanalyse darum in der Lösung des Problems einer geeigneten Faktorisierung von $\sum_{XX} - E^2 = AA^T$ in der Form $\sum_{XX} - E^2 = AA^T$ besteht. *Das Zustandekommen der Grundgleichung* $\sum_{XX} = AA^T + E^2$ *durch* *ein lineares Modell soll hier unerheblich sein.* Darüber hinaus halten wir fest, dass im Rahmen der Hauptkomponentenmethode keinerlei Einzelrestvarianzen vorkommen. Das Kommunalitätenproblem der Faktorenanalyse existiert für die meisten Anwendungen der Faktorenanalyse daher ohnehin nicht und die angestrebte Faktorisierung reduziert sich auf die Form $\sum_{XX} = AA^T$. Letzteres hat auch zur Folge, dass die Genauigkeit der Reproduktion von $\sum_{XX}$ durch AA^T jederzeit exakt kontrollierbar ist.

Konfrontiert mit dem Argument, dass es sich bei der den Probanden vorgelegten Bewertungsskala im Grunde um die Diskretisierung einer Intervallskala handelt, erfahren dann Methoden der Faktorenanalyse, die eigentlich nur für auf Intervallskalenniveau gemessene Daten geeignet scheinen, eine nachträgliche Berechtigung. Obwohl hier keine direkte Auseinandersetzung mit diesem Argument stattfinden wird, soll dennoch kurz erwähnt werden, dass dieses Argument problematisch ist, da den Probanden zur Beurteilung der ihnen vorgelegten Einflussgrößen keine Intervallskala, sondern ausdrücklich eine *endliche Ordinalskala* (also eine endliche Bewertungs- oder Ratingskala) vorgelegt wird, und sie dadurch gezwungen sind, diskret abzustimmen und sich deshalb *gerade nicht* auch für Zwischenwerte, also Werte innerhalb eines Intervalls entscheiden dürfen. Daher soll durch einen sehr allgemeinen Ansatz versucht werden zu klären, unter welchen Umständen die aus den beobachtbaren Daten gewonnene Korrelationsmatrix auch für nicht notwendig intervallskalierte Daten geeignet ist. Darüber hinaus ist dann natürlich noch die Frage zu diskutieren, unter welchen Bedingungen auch dieses vereinfachte faktorenanalytische Modell mit dem Skalenniveau (Messniveau) der beobachtbaren Daten verträglich sein kann, wenn dieses geringer ist als das Skalenniveau (Messniveau) einer Intervallskala. In den Teilen 3 und 4 dieses Abschnitts wird auf diese *Skalierungs- und Adäquatheitsprobleme der Faktorenanalyse* eingegangen.

4.2.1 Zur Kompatibilitätsproblematik der Faktorenanalyse

Erstes Anliegen ist es, auf folgende Problematik vieler Verfahren der Faktorenanalyse hinzuweisen, die *unabhängig vom Skalenniveau und von Wahrscheinlichkeitsverteilungen* immer dann auftreten kann, wenn die beobachtbaren Zufallsvariablen, die selbstverständlich alle den gleichen Definitionsbereich Ω haben sollen, auch einen als gleich zu interpretierenden Wertebereich (Messbereich) haben, der die Eigenschaft hat, dass reelle Zahlen m und t existieren können, so dass für mindestens eine Zufallsvariable X die transformierte Zufallsvariable $m \cdot X + t$ von X verschieden ist und ebenfalls in den gleichen Wertebereich (Messbereich) wie X fällt. Dabei setzen wir zunächst voraus, dass m größer als Null ist. Im Zusammenhang mit dem noch vorzustellenden Beispiel soll allerdings auch die Möglichkeit m kleiner Null erörtert werden. Diese Eigenschaft des Wertebereichs (Messbereichs) ist nur dann auszuschließen, wenn er *höchstens* zwei Werte umfasst. Will man nachfolgend beschriebene Problematik, die bereits einleitend als die *Kompatibilitätsproblematik zwischen beobachteten*

Daten einerseits und entsprechender Korrelationsmatrix andererseits (kurz: *Kompatibilitätsproblematik der Faktorenanalyse*) bezeichnet wurde, vermeiden, so muss, noch unscharf formuliert, der Wertebereich (Messbereich) der beobachtbaren Zufallsvariablen, da man andernfalls die beobachtbaren Zufallsvariablen nicht unterscheiden kann, *genau* zwei Werte enthalten. Im Folgenden wird der gemeinsame für alle beobachtbaren Zufallsvariablen als gleich zu interpretierende Wertebereich mit $\mathbb{W}$ bezeichnet. Grundsätzlich nehmen wir an, dass $\mathbb{W}$ eine Teilmenge von $\mathbb{R}$ ist. Im Folgenden sei mit $\mathfrak{X}$ die Menge derjenigen Zufallsvariablen bezeichnet, die zur Beschreibung des möglichen Probandenverhaltens zugelassen werden sollen. *Zur Vermeidung zusätzlicher - eher künstlicher - Komplikationen im Zusammenhang mit der Berechnung des Korrelationskoeffizienten setzen wir voraus, dass $\mathfrak{X}$ keine fast sicher konstanten Zufallsvariablen enthält.* Dann sei gefordert, dass $\mathfrak{X}$ folgende zwei Postulate erfüllt, wobei wir für jede Teilmenge M der Grundmenge Ω mit i_M die entsprechende Indikatorvariable kennzeichnen.

> **P1:** *Sei $\mathbb{V}= v_1,\ldots,v_n$ eine endliche Teilmenge von $\mathbb{W}$. Dann existieren paarweise disjunkte Teilmengen $V_1,\ldots,V_n$ von Ω, deren Wahrscheinlichkeiten alle größer als Null sind und für die $\sum_{k=1}^{n} v_k \cdot i_{V_k}$ in $\mathfrak{X}$ enthalten ist.*
>
> **P2:** *Seien $W_1,\ldots,W_n$ paarweise disjunkte Teilmengen von Ω und $w_1,\ldots,w_n$ reelle Zahlen, zu denen Zufallsvariablen aus $\mathfrak{X}$ existieren, die diese Werte mit von Null verschiedener Wahrscheinlichkeit annehmen. Dann gilt: $\sum_{k=1}^{n} w_k \cdot i_{W_k} \in \mathfrak{X}$.*

Zum Verständnis der Postulate **P1** und **P2** denke man z.B. an einen Test oder eine Bewertungsskala (siehe nachfolgendes Beispiel), der oder die einer Probandengruppe vorgelegt wurde (Ω ist in diesem Fall die Menge aller möglichen Probanden, denen der Test oder die Bewertungsskala vorgelegt werden kann).

Postulat **P1** besagt dann: Wenn $v_1,\ldots,v_n$ mögliche Punktzahlen oder Bewertungen sind, dann kann es auch Probanden geben, die diese Punktzahlen erreichen oder diese Bewertungen vornehmen.

Postulat **P2** beinhaltet darüber hinaus: Wenn $W_1,\ldots,W_n$ irgendwelche paarweise disjunkten Mengen von Probanden sind und $w_1,\ldots,w_n$ mögliche Punktzahlen oder Bewertungen, so darf nicht a priori ausgeschlossen werden, dass für alle $1 \le k \le n$ die in der Menge W_k zusammengefassten Probanden die Punktzahl w_k bekommen oder die Bewertung w_k vorgenommen haben.

In jeder konkreten Untersuchungssituation muss das dieser Untersuchungssituation zugrunde liegende wahrscheinlichkeitstheoretische Modell also die Postulate **P1** und **P2** ganz selbstverständlich erfüllen.

Im Kapitel über die Faktorenanalyse schreibt Bortz (1993) in seinem weitverbreiteten Standardwerk „Statistik für Sozialwissenschaftler" nach Vorstellung eines Eingangsbeispiels zweierlei. Zunächst formuliert er:

> *„Das Beispiel verdeutlicht die erste wichtige Eigenschaft der Faktorenanalyse. Sie ermöglicht es, ohne entscheidenden Informationsverlust viele wechselseitig mehr oder wenig hoch korrelierende Variablen durch wenige voneinander unabhängige Faktoren zu ersetzen"* (ebd., S. 617).

Wenig später liest man allerdings:

> *„Die Faktorenanalyse liefert jedoch keinerlei Anhaltspunkte dafür, was das Gemeinsame dieser Fragen ist, sondern lediglich, dass die untersuchte Stichprobe diese Fragen sehr ähnlich beantwortet hat"* (ebd., S. 617).

Im Folgenden wird versucht zu erörtern, dass Aussagen wie *„ohne entscheidenden Informationsverlust"* und *„sehr ähnlich"* durchaus einer näheren Reflexion bedürfen. Diese Aussagen erlauben nämlich einen gewaltigen Interpretationsspielraum, da die Begriffe *Informationsverlust* und *ähnlich* durch den jeweiligen Anwender individuell vorgeprägt sein können.

Ein erstes erläuterndes Beispiel

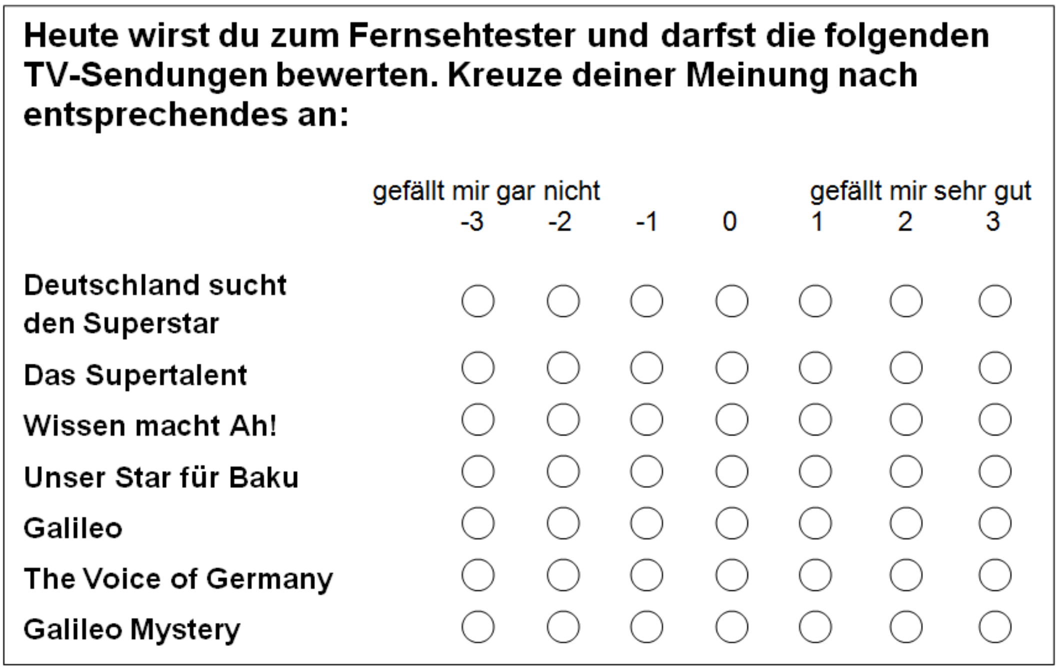
Im folgenden Beispiel, das auch in der empirischen Sozialforschung vorkommen kann (in diesem Zusammenhang schreibt Bortz (1993, S. 619):

> *„Die Entwicklung der Faktorenanalyse wäre zweifellos nicht so stürmisch verlaufen, wenn nicht gleichzeitig insbesondere von Psychologen die herausragende Bedeutung dieses Verfahrens für sozialwissenschaftliche Fragestellungen erkannt und immer wieder nach differenzierteren und mathematisch besser abgesicherten Analysemöglichkeiten verlangt worden wäre."*),

gehen wir von der Bewertungsskala -3, -2, -1, 0, 1, 2, 3 als dem gemeinsamen Wertebereich (Messbereich) von sieben Zufallsvariablen $X_1, X_2, ..., X_7$ aus. Selbstverständlich wäre auch jede andere mindestens drei unterschiedliche Bewertungen ermöglichende Bewertungsskala denkbar. Darüber hinaus betrachten wir der Einfachheit halber acht Probanden $P_1, P_2, ..., P_8$ (die Zahlen „7" und „8" sind zufällig gewählt und für die folgende Betrachtung völlig irrelevant), die den Definitionsbereich der Zufallsvariablen $X_1, X_2, ..., X_n$ repräsentieren. $X_i(P_k)$ $(1 \leq i \leq 7, 1 \leq k \leq k)$ entspricht dann der Bewertung des Probanden P_k im Hinblick auf die Zufallsvariable X_i. Ein so kleines Beispiel entspricht ganz sicher nicht der Praxis empirischer Forschungen, verdeutlicht jedoch das uns interessierende Phänomen, welches bei großen Datenmengen oft verdeckt bleibt.

Zum besseren Verständnis der folgenden Ausführungen stelle man sich vor, dass es sich wie in dieser Arbeit um die Urdaten einer Schülerbefragung zur Ursachenforschung des Fernsehverhaltens handelt.

Ziel des Fragebogens ist es festzustellen, welche Faktoren das Fernsehverhalten der Schülerinnen und Schüler bestimmen. Der Fragebogen könnte also wie folgt aussehen:


Heute wirst du zum Fernsehtester und darfst die folgenden TV-Sendungen bewerten. Kreuze deiner Meinung nach entsprechendes an:

	gefällt mir gar nicht -3	-2	-1	0	1	2	gefällt mir sehr gut 3
Deutschland sucht den Superstar	○	○	○	○	○	○	○
Das Supertalent	○	○	○	○	○	○	○
Wissen macht Ah!	○	○	○	○	○	○	○
Unser Star für Baku	○	○	○	○	○	○	○
Galileo	○	○	○	○	○	○	○
The Voice of Germany	○	○	○	○	○	○	○
Galileo Mystery	○	○	○	○	○	○	○

Abbildung 4.1 Fragebogen für das erste Beispiel

Zunächst wird davon ausgegangen, dass dieser Fragebogen einer Gruppe von Schülerinnen und Schülern vorgelegt wurde, die die sieben Sendungen alle schon häufiger gesehen haben.

Im Folgenden kürzen wir die obigen Einflussfaktoren repräsentierenden Zufallsvariablen nacheinander mit X_1 (Deutschland sucht den Superstar), X_2 (Das Supertalent), X_3 (Wissen macht Ah!), X_4 (Unser Star für Baku), X_5 (Galileo), X_6 (The Voice of Germany) und X_7 (Galileo Mystery) ab. Dann nehmen wir an, dass das Ergebnis dieser Befragung in nachfolgender Datenmatrix **D1** erfasst ist.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	3	2	3	3	2	2	3	3
X_2	2	3	3	3	2	2	3	3
X_3	-3	-3	-3	-2	-3	-3	-2	-2
X_4	2	2	3	2	2	2	3	3
X_5	-3	-3	-1	-1	-3	-1	-1	-1
X_6	2	2	3	3	2	2	2	3
X_7	-3	-3	-2	-2	-3	-3	-2	-3

Abbildung 4.2 Datenmatrix D1

Die Datenmatrix **D1** legt offensichtlich die Interpretation nahe, dass einerseits die Zufallsvariablen X_1, X_2, X_4 und X_6 und andererseits die Zufallsvariablen X_3, X_5 und X_7 auf je einen gemeinsamen ihnen zugrunde liegenden Faktor hindeuten. Vergleichen wir nämlich die Zeilen der Datenmatrix **D1** miteinander, so ist für die durch die Variablen X_1, X_2, X_4 und X_6 repräsentierten Zeilen eine recht hohe Ähnlichkeit im Abstimmungsverhalten der Probanden ablesbar. Das Gleiche trifft (mit kleinen Abstrichen) auch für die durch die Variablen X_3, X_5 und X_7 repräsentierten Zeilen zu. Vergleicht man jedoch je eine Zeile der durch die Variablengruppe X_1, X_2, X_4 und X_6 gegebene Zeile mit je einer Zeile der durch die Variablengruppe X_3, X_5 und X_7 gegebenen Zeilen, so muss man das entsprechende *Abstimmungsverhalten* der Probanden in diesen Zeilen wohl als unterschiedlich ansehen. Dabei deutet der durch die Zufallsvariablen X_1, X_2, X_4 und X_6 erfasste Faktor auf reines Unterhaltungsfernsehen hin, während der zweite Faktor eher auf wissensfördernde Sendungen hinweist. Der Einfluss des ersten Faktors ist als hoch und der des zweiten Faktors eher als niedrig einzuschätzen.

Ohne irgendwelche Zusatzannahmen ist eine Interdependenz der vermuteten Faktoren in die vorgestellten Daten eigentlich nicht hinein interpretierbar. *Letzteres wird auch bestätigt, wenn man, wie in empirischen Untersuchungen nicht unüblich, Verfahren der Faktorenanalyse durch Methoden der Clusteranalyse oder Multidimensionalen Skalierung, die auf dem Euklidischen Abstand, der City-Block-Metrik oder einem verwandten Distanzmaß basieren, ergänzen würde.* Die Kompatibilität verschiedener datenanalytischer Verfahren, auf die in Abschnitt 4 noch einmal kurz eingegangen wird, ist sicher wünschenswert.

Nochmals zurück zu dem Ausgangsbeispiel: Wir stellen uns vor, dass im Rahmen der Untersuchung noch eine zweite Gruppe von Schülerinnen und Schülern befragt wurde, welche das Fernsehprogramm weniger intensiv verfolgt. Nun unterstellen wir der Gruppe, dass sie beide oben erörterten Einflussfaktoren als sehenswert einschätzen. Die entsprechende Datenmatrix **D2** könnte dann wie folgt aussehen:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	3	2	3	3	2	2	3	3
X_2	2	3	3	3	2	2	3	3
X_3	2	2	2	3	2	2	3	3
X_4	2	2	3	2	2	2	3	3
X_5	2	2	3	3	2	3	3	3
X_6	2	2	3	3	2	2	2	3
X_7	2	2	3	3	2	2	3	2

Abbildung 4.3 Datenmatrix D2

Entsprechend sollte ein angemessenes faktorenanalytisches Verfahren nur einen Faktor extrahieren.

Man kann sich jedoch schnell davon überzeugen, dass für die in der Datenmatrix **D1** erfassten Werte der beobachteten Zufallsvariablen gilt:

$$X_3 = 2 \cdot (2,2,2,3,2,2,3) - 5\,,$$
$$X_5 = 2 \cdot (2,2,3,3,2,3,3,3) - 5 \text{ und}$$
$$X_7 = 2 \cdot (2,2,3,3,2,2,3,2) - 5\,.$$

Dies bedeutet, dass die Datenmatrix **D2** zur gleichen unten aufgeführten Korrelationsmatrix **Corr1** (Kovarianzmatrix der standardisierten Zufallsvariablen) führt, wie die Datenmatrix **D1**. Die Korrelationsmatrix interpretiert nämlich in jeder Zeile der beobachteten Datenmatrix den Wert x in der Relation $x < y$ als negatives Abstimmungsverhalten der Probanden, während sie den Wert y als positives Abstimmungsverhalten der Probanden ansieht.

In beiden Fällen liefert die Faktorenanalyse also das gleiche Ergebnis. *Ein Anwender faktorenanalytischer Methoden ließe sich (in diesem Beispiel) daher durch die Methode vorschreiben, wie sie oder er die beobachteten Daten zu interpretieren hat.*

Dieses Phänomen wird besonders deutlich, wenn wir, wie wir sofort verifizieren, berücksichtigen, dass sowohl **D1** *als auch* **D2** *zur gleichen Datenmatrix* **DS** *der standardisierten Zufallsvariablen führen.*

cor	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1						
X_2	0,467	1					
X_3	0,6	0,6	1				
X_4	0,6	0,6	0,467	1			
X_5	0,467	0,467	0,6	0,6	1		
X_6	0,6	0,6	0,467	0,467	0,6	1	
X_7	0,6	0,6	0,467	0,467	0,6	0,467	1

Abbildung 4.4 Korrelationsmatrix DS

Die bisherige Erörterung ist unabhängig von dem aufgrund der Thematik der Arbeit gewählten Kontext *Ursachen für Fernsehkonsum*. Ebenso könnten die Zufallsvariablen anders interpretiert werden, wie die folgenden beiden Beispiele demonstrieren:

> Die Zufallsvariablen $X_1, X_2, X_3, X_4, X_5, X_6$ und X_7 könnten - wie später in dieser Arbeit noch tatsächlich erhoben - Fähigkeiten beschreiben, die einen guten Mathematiker ausmachen. Man denke z.B. bei X_1 an die Fähigkeit „logisches Denkvermögen", bei X_2 an die Fähigkeit „räumliche Vorstellungskraft", bei X_3 an die mögliche Fähigkeit „Fantasie", bei X_4 an die Fähigkeit „Ausdauer und Fleiß", bei X_5 an die Fähigkeit „Sprachvermögen", bei X_6 an die Fähigkeit „Abstraktionsvermögen" und bei X_7 an die Fähigkeit „Gedächtnisleistung". Hier bedeutet „-3", dass die entsprechende Fähigkeit keinen Einfluss auf die Güte eines guten Mathematikers besitzt, während „3" bedeutet, dass der Einfluss der entsprechenden Fähigkeit als äußerst wichtig eingeschätzt wird. Tatsächlich ist der Versuch der Reduzierung von Einflussgrößen auf wenige, diese Einflussfaktoren erklärende, Faktoren nicht untypisch für die Anwendung faktorenanalytischer Verfahren (vgl. z.B. den in KNOCHE (1990) vorgestellten Wilde Test zur Untersuchung von Niveau und Struktur intellektueller Fähigkeiten).

> Die Zufallsvariablen $X_1, X_2, X_3, X_4, X_5, X_6$ und X_7 können - im Rahmen einer entsprechenden Befragung Jugendlicher - aber auch Einflussgrößen für Zukunftsängste von jungen Menschen beschreiben. Man denke z.B. bei X_1 an die Einflussgröße „wirtschaftliche Entwicklung", bei X_2 an die Einflussgröße „Umweltgefahren", bei X_3 an die Einflussgröße „berufliche Karrieremöglichkeiten", bei X_4 an die Einflussgröße „Terrorismusgefahr", bei X_5 an die Einflussgröße „privates Glück", bei X_6 an die Einflussgröße „politische Entwicklung" und bei X_7 an die

Einflussgröße „Gesundheit". Hier bedeutet „-3", dass keine nennenswerte Angst und „3", dass eine große Angst besteht.

Jede Faktorenanalyse (ob Hauptkomponentenmethode, Hauptfaktorenanalyse, Minres-Faktorenanalyse oder ML Faktorenanalyse usw.) bestätigt die Vermutung des einen gemeinsamen Faktors. Darüber hinaus fällt auf, dass die Korrelationen zwischen den Zufallsvariablen recht niedrig sind, und das, obwohl sich je zwei in der Datenmatrix **D2** erfasste Zufallsvariablen um nur genau zwei Abweichungen voneinander unterscheiden. Dieser Eindruck verwischt sich jedoch, wenn man bedenkt, dass $cor(X_i, X_j) = cor(X_i, X_j^{'})$ für alle $i \in \{1,2,4,6\}$ und alle $j \in \{3,5,7\}$ gilt. Im Übrigen ist für je zwei in der Datenmatrix **D2** erfasste Zufallsvariable die Korrelation 0,467, wenn ihre Erwartungswerte gleich sind, andernfalls beträgt sie 0,6.

Diese Beobachtungen verdeutlichen noch einmal, dass ***die inhaltliche Ähnlichkeit*** *im Abstimmungsverhalten der Probanden und die durch den* ***Korrelationskoeffizienten festgelegte Ähnlichkeit*** *im Abstimmungsverhalten der Probanden auseinanderklaffen.*

Diese Widersprüche zwischen inhaltlicher Bedeutung der Messergebnisse einerseits und Rechenergebnis einer Faktorenanalyse andererseits lässt das Ergebnis der Faktorenanalyse - ohne weitere Analysen der Datenstrukturen - kaum interpretierbar werden (da die Überlegungen metrische und normalverteilte Daten nicht ausgeschlossen haben, auch dann noch, wenn metrische und normalverteilte Daten beobachtet wurden).

Der Korrelationskoeffizient als Maß für die Interdependenz der Zufallsvariablen ist in obigem Beispiel daher (scheinbar) nicht geeignet. Diese (scheinbare) Nichteignung des Korrelationskoeffizienten liegt in seinem durch ihn verursachten erheblichen ***Informationsverlust*** *begründet, denn offensichtlich beinhaltet die Datenmatrix* **D1** *völlig andere Informationen als die Datenmatrix* **D2** *und dennoch führen beide zu der gleichen Korrelationsmatrix, dem Ausgangspunkt einer Faktorenanalyse.*

Tatsächlich impliziert die Tatsache, dass durch je zwei Punkte des $\mathbb{R}^2$ *genau eine Gerade verläuft, dass eine Datenmatrix* $\mathbf{D}$, *deren Werte auf der Bewertungsskala* -3, -2, -1, 0, 1, 2, 3 *liegen, genau dann die gleiche Matrix der standardisierten Variablen wie* **D1** *oder* $(-1)\cdot\mathbf{D1}$ *bestimmt, und somit die gleiche Korrelationsmatrix wie* **D1** *realisiert, wenn für alle* $1 \le k \le 7$ *streng monoton steigende oder für alle* $1 \le k \le 7$ *streng monoton fallende Abbildungen* $\varphi_1, \ldots \varphi_7 : (\mathbb{R}, \le) \mapsto (\mathbb{R}, \le)$ *existieren mit*

$\varphi_1(2) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_1(3) \in \{-3,-2,-1,0,1,2,3\}$,

$\varphi_2(2) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_2(3) \in \{-3,-2,-1,0,1,2,3\}$,

$\varphi_3(-1) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_3(1) \in \{-3,-2,-1,0,1,2,3\}$,

$\varphi_4(2) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_4(3) \in \{-3,-2,-1,0,1,2,3\}$,

$\varphi_5(-1) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_5(1) \in \{-3,-2,-1,0,1,2,3\}$,

$\varphi_6(2) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_6(3) \in \{-3,-2,-1,0,1,2,3\}$ *und*

$\varphi_7(-1) \in \{-3,-2,-1,0,1,2,3\}$ *und* $\varphi_7(1) \in \{-3,-2,-1,0,1,2,3\}$, *so dass* $\boldsymbol{D}$ *die Form*

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	$\varphi_1(3)$	$\varphi_1(2)$	$\varphi_1(3)$	$\varphi_1(3)$	$\varphi_1(2)$	$\varphi_1(2)$	$\varphi_1(3)$	$\varphi_1(3)$
X_2	$\varphi_2(2)$	$\varphi_2(3)$	$\varphi_2(3)$	$\varphi_2(3)$	$\varphi_2(2)$	$\varphi_2(2)$	$\varphi_2(3)$	$\varphi_2(3)$
X_3	$\varphi_3(-3)$	$\varphi_3(-3)$	$\varphi_3(-3)$	$\varphi_3(-2)$	$\varphi_3(-3)$	$\varphi_3(-3)$	$\varphi_3(-2)$	$\varphi_3(-2)$
X_4	$\varphi_4(2)$	$\varphi_4(2)$	$\varphi_4(3)$	$\varphi_4(2)$	$\varphi_4(2)$	$\varphi_4(2)$	$\varphi_4(3)$	$\varphi_4(3)$
X_5	$\varphi_5(-3)$	$\varphi_5(-3)$	$\varphi_5(-2)$	$\varphi_5(-2)$	$\varphi_5(-3)$	$\varphi_5(-2)$	$\varphi_5(-2)$	$\varphi_5(-2)$
X_6	$\varphi_6(2)$	$\varphi_6(2)$	$\varphi_6(3)$	$\varphi_6(3)$	$\varphi_6(2)$	$\varphi_6(2)$	$\varphi_6(2)$	$\varphi_6(3)$
X_7	$\varphi_7(-3)$	$\varphi_7(-3)$	$\varphi_7(-2)$	$\varphi_7(-2)$	$\varphi_7(-3)$	$\varphi_7(-3)$	$\varphi_7(-2)$	$\varphi_7(-3)$

Abbildung 4.5 Tabelle D

hat. Diese Beobachtung, die im Weiteren mit (*) abgekürzt wird, verdeutlicht das Ausmaß eines (möglichen) Informationsverlusts, der beim Übergang zur Korrelationsmatrix entstehen kann. In vorgestelltem Beispiel rechnet man nämlich nach, dass $\mathbf{D}$ genau $2 \cdot \binom{7}{2}^7 = 2 \cdot 21^7$ $= 3602177082$, also mehr als 3,6 Milliarden Datenmatrizen möglich sind, die alle entweder zur gleichen Matrix der standardisierten Variablen wie **D1** oder $(-1)\cdot\mathbf{D1}$ und damit zur gleichen Korrelationsmatrix führen.

Und jede dieser Datenmatrizen $\mathbf{D}$ *bedeutet eine andere Bewertung durch die Probanden, also möglicherweise auch eine andere Interpretation der Daten durch den Anwender.*

In diesem Zusammenhang sei es zudem erlaubt auf die beiden folgenden Aspekte, die im Zusammenhang mit der durch einen Informationsverlust entstehenden Kompatibilitätsproblematik der Faktorenanalyse besonders auffallen, hinzuweisen:

> **Aspekt 1:** *Bei größeren Datenmengen ist die Kompatibilitätsproblematik oft nicht sichtbar.*

Aspekt 2: *Faktorenanalytische (datenanalytische) Verfahren werden oft angewandt, weil sie sich nach der festen Überzeugung des Benutzers in der Vergangenheit bei ähnlich gelagerten Anwendungen bewährt haben. Dies bedeutet, dass die erhaltenen Ergebnisse und ihre Interpretation zwar kompatibel mit vergleichbaren früheren Untersuchungen sind, aber nicht unbedingt mit den tatsächlich beobachteten Daten.*

Da Anwender Beispielen gegenüber (oft auch zu recht) häufig den Verdacht der Künstlichkeit - und damit einhergehender geringer Praxisrelevanz - äußern, sei noch einmal betont, dass die Daten in obigem Beispiel nur eine untergeordnete Rolle spielen. Das in dem angegebenen Beispiel zum Ausdruck kommende Phänomen beruht einzig und allein auf der Tatsache, dass der Korrelationskoeffizient gegenüber affinen Transformationen invariant ist und ist daher, wenn auch vom Anwender meist unbemerkt, in allen beobachtbaren Datenmatrizen, die mindestens eine Zeile **Z** *enthalten, die folgende Bedingungen* ***nicht*** *erfüllen, versteckt:*

Z1: *Sowohl die kleinstmögliche Punktzahl (Bewertung) „k" als auch die größtmögliche Punktzahl „g" ist in* **Z** *enthalten.*

Z2: *Es existieren in* **Z** *zwischen „k" und „g" liegende Punktzahlen (Bewertungen)* p_i $(i \in I)$, *zu denen keine Punktzahlen (Bewertungen)* $k' < p_i' < g'$ $(i \in I)$ *existieren können, so dass die Paare* (k, k'), (p_i, p_i') *und* (g, g') *auf einer (gemeinsamen) Geraden liegen.*

Tatsächlich gilt Beobachtung (*) für jede beliebige Datenmatrix, wenn man die entweder stets streng monoton steigenden oder stets streng monoton fallenden Abbildungen φ_k $(1 \le k \le 7)$ durch beliebige entweder streng monoton steigende oder streng monoton fallende affine Transformationen ersetzt. Wir halten darum genauer fest: Seien **D** eine beliebige aber fest vorgegebene Datenmatrix und **C** die ihr entsprechende Korrelationsmatrix. Dann bezeichnen wir mit $\mathcal{D}$ die Menge aller Datenmatrizen, die auf der gleichen Punktevergabe (Bewertung) wie **D** beruhen, und deren entsprechende Korrelationsmatrix **C** ist. Nun sehen wir zwei Datenmatrizen $D' \in \mathcal{D}$ und $D'' \in \mathcal{D}$ als äquivalent an, wenn für die ihnen entsprechenden Matrizen **Ds'** bzw. **D**s'' der standardisierten Variablen entweder **Ds'=D**s'' oder **Ds'=**$(-1)\cdot$**D**s'' gilt. Letzteres ist dazu äquivalent, dass zu jeder Zeile **Z'** von **D'** eine affine Transformation der Form $m_{Z'} \cdot x + t_{Z'}$ existiert, die **Z'** in die entsprechende Zeile **Z''** von **D''** überführt, wobei entweder $m_{Z'} > 0$ für alle Zeilen **Z'** von **D'** oder $m_{Z'} < 0$ für alle Zeilen **Z'** von **D'** gilt. Die Anzahl der auf diese Weise definierten Äquivalenzklassen können wir **nicht** berechnen, wohl aber die Anzahl der Datenmatrizen in einer Äquivalenzklasse. Sei dazu „n" die Anzahl der Zeilen von **D** und „v" die vorgegebene Anzahl möglicher Punkte oder Bewertungen. Dann besteht die Äquivalenzklasse **[D]** von **D** aus genau einer Datenmatrix, wenn jede Zeile **Z** von **D** beide Bedingungen **Z1** und **Z2** erfüllt und aus höchstens $2 \cdot \binom{v}{2}^n$ verschiedenen Datenmatrizen, nämlich, wie man sich sofort verdeutlicht, genau dann, wenn jede Zeile **Z** von **D** genau zwei (verschiedene) Werte enthält. Die beobachteten Datenmengen können dabei völlig unterschiedlicher Natur sein.

Bei naturwissenschaftlichen Messungen werden in der Regel verschieden geartete Größen, wie Blutzuckergehalt, Gewicht, Körpergröße, Alter, Blutdruck usw. im Hinblick auf ihre möglichen Interdependenzen (gemeinsame latente Faktoren) hin untersucht. Hier haben die gemessenen Daten keinen als gemeinsam zu interpretierenden Wertebereich (Messbereich) und die beschriebene Kompatibilitätsproblematik der Faktorenanalyse wird kaum auftreten können, wobei sie nicht völlig auszuschließen ist.

Wir halten daher, zunächst noch recht vage formuliert (die genauere Untersuchung erfolgt im 4. Abschnitt), fest, dass die beschriebene Kompatibilitätsproblematik der Faktorenanalyse bei verschiedenen (unterschiedlich zu interpretierenden) Wertebereichen (Messbereichen) der Zufallsvariablen oder wenn der Wertebereich (Messbereich) der Zufallsgrößen nur zwei Werte umfasst wohl nicht auftreten wird.

Das Überprüfen der inhaltlichen Kompatibilitätsproblematik von Datenmatrix einerseits und Korrelationsmatrix (Kovarianzmatrix der standardisierten Zufallsvariablen) andererseits sollte stets vor Beginn des Anwendens eines faktorenanalytischen Verfahrens erfolgen, um die Angemessenheit eines faktorenanalytischen Verfahrens abzusichern.

Ein weiteres Beispiel

Folgendes weitere Beispiel, welches das erste Beispiel dahin gehend ergänzt, dass jede der bekannten Methoden der Faktorenanalyse die Anzahl der Faktoren überschätzen würde, dient dazu, das Problem der (inhaltlichen) Kompatibilität faktorenanalytischer Methoden besonders deutlich sichtbar werden zu lassen. Dazu betrachten wir je die drei Tests A_1, A_2 und A_3 bzw. L_1, L_2 und L_3 zur „Analysis" bzw. zur „Linearen Algebra", die 6 Probanden $P_1, \ldots, P_6$ vorgelegt bekommen haben. Zu erreichen waren jeweils „0" bis maximal „100" Punkte (vgl. HARTUNG und ELPELT (1984, S. 523)). Die Testergebnisse seien in folgender Datenmatrix **D1*** zusammengefasst, wobei die Datenmatrix **D1*** nur eine Interpretation zulässt, nämlich die, dass die Testergebnisse ganz ausgezeichnet sind, dass alle Probanden sowohl die Tests zur „Analysis" als auch die Tests zur „Linearen Algebra" nahezu perfekt lösen konnten. Es kann also nur einen zugrunde liegenden Faktor geben, der auf ununterscheidbare (hervorragende) Kenntnisse sowohl in „Analysis" als auch in „Linearer Algebra" hinweist.

	P_1	P_2	P_3	P_4	P_5	P_6
A_1	100	99	99	100	99	100
L_1	99	100	100	99	100	99
A_2	100	99	99	100	100	100
L_2	99	100	100	99	100	99
A_3	100	99	99	99	99	100
L_3	99	100	99	99	100	99

Abbildung 4.6 Datenmatrix D1*

Was aber geschieht, wenn die Datenmatrix **D1*** mit Hilfe irgendeiner Methode der Faktorenanalyse, der die resultierende (empirische) Korrelationsmatrix zugrunde liegt, ausgewertet wird? Durch jede dieser Methoden wird die Punktzahl „99" verfälscht zum „Nicht lösen können der Probanden des entsprechenden Tests", während die Punktzahl „100" als „Lösen können der Probanden des entsprechenden Tests" interpretiert wird. Codieren wir daher „Nicht lösen können der Probanden des entsprechenden Tests" mit „0" und „Lösen Können der Probanden des entsprechenden Tests" mit „1", so liest jede Methode der Faktorenanalyse, der die (empirische) Korrelationsmatrix zugrunde liegt, die Datenmatrix **D1*** wie folgt:

	P_1	P_2	P_3	P_4	P_5	P_6
A_1	1	0	0	1	0	1
L_1	0	1	1	0	1	0
A_2	1	0	0	1	1	1
L_2	0	1	1	0	1	0
A_3	1	0	0	0	0	1
L_3	0	1	0	0	1	0

Abbildung 4.7 Datenmatrix D1* nach Transformation

Leserinnen bzw. Leser, die geneigt sind, die Künstlichkeit des Beispiels zu kritisieren, was die grundsätzliche in diesem Abschnitt besprochene Problematik faktorenanalytischer Verfahren ohnehin nicht berühren würde, seien darauf hingewiesen, dass es jederzeit möglich ist, durch Transformationen der Daten mit Hilfe stets streng monoton steigender oder stets streng monoton fallender (reellwertiger) Abbildungen φ_k $\left(1 \le k \le 6\right)$ der Datenmatrix **D1*** ein sehr realistisches Aussehen zu geben (vgl. die unter Beobachtung (*) vorgenommenen Hinweise). In diesem Zusammenhang vergewissere sich die Leserin bzw. der Leser der Tatsache, dass die erhaltene „0-1-Datenmatrix" durch die affine Transformation $x \mapsto x - 99$ aus der Datenmatrix **D1*** hervorgeht.

Obwohl im Grunde fast unnötig sei noch die der Datenmatrix **D1*** entsprechende Korrelationsmatrix **Corr1*** angegeben:

cor	A_1	L_1	A_2	L_2	A_3	L_3
A_1	1					
L_1	-1	1				
A_2	0,707	-0,707	1			
L_2	-1	1	-0,707	1		
A_3	0,707	-0,707	0,5	-0,707	1	
L_3	-0,707	0,707	-0,25	0,707	-0,5	1

Abbildung 4.8 Korrelationsmatrix der Datenmatrix D1*

Es ist erkennbar, dass die Tests zur „Analysis" einerseits und die Tests zur „Linearen Algebra" andererseits recht hoch miteinander korrelieren, während die Tests zur „Analysis" und die Tests zur „Linearen Algebra" untereinander ausschließlich negative Korrelationen aufweisen. Jedes Verfahren der Faktorenanalyse, dessen Ausgangspunkt die (empirische) Korrelationsmatrix ist, würde also die beiden sich *gegenseitig nicht beeinflussenden* Faktoren *Analysisfertigkeiten* und *Fertigkeiten in Linearer Algebra* extrahieren.

Derartige Ungereimtheiten können leider nicht ausgeschlossen werden, wenn man sich als Anwender von der Methode vorschreiben lässt, wie die beobachteten Daten zu interpretieren sind. Daher sei darüber hinaus ebenfalls erlaubt, all diejenigen, die den Erfolg eines datenanalytischen Verfahrens auch an der Interpretierbarkeit des Ergebnisses messen, darauf hinzuweisen, dass das in obigem Beispiel erhaltene Ergebnis ganz hervorragend interpretierbar, aber leider völlig falsch ist.

Das Hauptziel dieses Abschnitts muss somit in der Entwicklung eines Vorschlags (Verfahrens) bestehen, welcher (welches) hilft, die beobachtbaren Ausgangsdaten ohne Informationsverlust so aufzubereiten bzw. darzustellen, dass eine Fehlinterpretation der Ausgangsdaten durch die (empirische) Korrelationsmatrix vermieden werden kann.

Bevor jedoch eine mögliche Lösung der Kompatibilitätsproblematik der Faktorenanalyse diskutiert wird, soll noch kurz auf das für die empirische Forschung in der Didaktik der Mathematik grundsätzliche Problem der *Adäquatheit* von Methoden der Faktorenanalyse im Bezug auf den Messbarkeitsgrad der beobachtbaren Daten (kurz: der *Skalierungsproblematik der Faktorenanalyse*) und das Problem der *Adäquatheit* des der Faktorenanalyse zugrunde liegenden Modells (kurz: *der Methodenproblematik der Faktorenanalyse*) eingegangen werden.

4.2.2 Zur Skalierungsproblematik der Faktorenanalyse und ihrer möglichen Behandlung

Auf den ersten Blick könnte man meinen, dass Methoden der Faktorenanalyse lediglich auf intervallskalierte Zufallsvariable (Daten) anwendbar sind, da der Korrelationskoeffizient, die Grundlage der Faktorenanalyse, lediglich gegenüber Skalentransformationen der Form $m \cdot x + t \; \left(m \in \mathbb{R} \setminus \{0\}, t \in \mathbb{R}\right)$ invariant ist.

Zur Präzisierung der *Skalierungsproblematik* der Faktorenanalyse betrachten wir daher eine Gruppe $\left(\mathbf{T}, \circ\right)$ von Transformationen $T : \mathbb{R} \mapsto \mathbb{R}$. Da die Zeilen der zugrunde liegenden Datenmatrix mit *endlichen* Zufallsvariablen identifiziert werden können, lässt sich die *Skalierungsproblematik* der Faktorenanalyse durch folgende Frage erfassen:

Welche Gestalt muss $\mathbb{W}$ *haben, so dass für alle Transformationen* $T \in \mathbf{T}$ *und alle endlichen Zufallsvariablen* $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ *die Invarianz* $cor(X,Y) = cor(T \circ X, T \circ Y)$ *gilt?*

In Ergänzung zu dieser Frage definiere man sehr allgemein, dass der Korrelationskoeffizient *skalierungsunabhängig* ist, wenn für alle bijektiven Abbildungen $B : \mathbb{R} \mapsto \mathbb{R}$ und alle endlichen Zufallsvariablen $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ die Invarianz $cor(X,Y) = cor(B \circ X, B \circ Y)$ gilt. Da sich die Skalierungsunabhängigkeit von Korrelationskoeffizienten leicht charakterisieren lässt, zitieren wir den entsprechenden Charakterisierungssatz, dessen Beweis im Anhang nachgelesen werden kann. In diesem Satz ist mit $Aut(\mathbb{R}, \leq)$ die Gruppe der Ordnungsautomorphismen auf $\left(\mathbb{R}, \leq\right)$ bezeichnet.

Satz 4.1:

Folgende Aussagen sind äquivalent:

(i) *Der Korrelationskoeffizient ist skalierungsunabhängig.*

(ii) *Für alle Ordnungsautomorphismen* $T \in Aut(\mathbb{R}, \leq)$ *und alle Zufallsvariablen* $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ *gilt die Gleichheit von* $cor(X,Y)$ *und* $cor(T \circ X, T \circ Y)$.

(iii) $\mathbb{W}$ *enthält höchstens zwei Elemente.*

Bei dichotomen Zufallsvariablen tritt die Skalierungsproblematik der Faktorenanalyse also gar nicht erst auf.

Ein mit der Skalierungsunabhängigkeit verwandter, aber wesentlich schwächerer Begriff, ist der Begriff der *Permutationsunabhängigkeit.* In der Clusteranalyse wird er z.B. von Jardine und Sibson (1971) im Zusammenhang mit dem Single-Linkage Verfahren gefordert. Im Rahmen des hier vorgestellten Zugangs zur Faktorenanalyse lässt er sich sehr allgemein wie folgt fassen:

Eine Abbildung $s : \mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ heißt *permutationsunabhängig*, wenn für alle endlichen Zufallsvariablen $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ und alle bijektiven Abbildungen $B : \mathbb{W} \mapsto \mathbb{W}$ stets $s(X,Y) = s(B \circ X, B \circ Y)$ gilt.

In konkreten Untersuchungssituationen sehen wir das Ziehen von Probanden, Testaufgaben usw. (kurz: Elementarereignissen) als gleich wahrscheinlich an. Deshalb ist folgende einfache Beobachtung vielleicht nicht ganz uninteressant:

Seien $X := \sum_{k=1}^{n} x_k \cdot i_{V_k}$ und $Y := \sum_{k=1}^{n} y_k \cdot i_{W_k}$ endliche Zufallsvariablen mit $x_i \neq x_j$ für alle $1 \leq i \neq j \leq n$ und $y_i \neq y_j$ für alle $1 \leq i \neq j \leq n$ und darüber hinaus $P(V_i) = P(V_j) = P(W_i) = P(W_j)$ für alle $1 \leq i, j \leq n$, Dann gilt für jede bijektive Abbildung $B : \mathbb{W} \mapsto \mathbb{W}$ die Gleichung $cor(X,Y) = cor(B \circ X, B \circ Y)$.

Der Zusammenhang zwischen Skalierungsunabhängigkeit und Permutationsunabhängigkeit wird in den Sätzen 2, 3 und 4 von Abschnitt 4 verdeutlicht.

4.2.3 Zur Modellproblematik der Faktorenanalyse und ihrer möglichen Behandlung

Eine Mindestanforderung, die an ein auf ein datenanalytisches Modell zurückgehendes Verfahren gestellt werden muss, ist die Forderung nach der *Adäquatheit* dieses Verfahrens. Zur ersten Annäherung an die Forderung nach der Adäquatheit eines Verfahrens betrachten wir ein beliebiges zur Bestimmung optimaler Lösungen geeignetes Verfahren V , das auf eine beobachtete Datenmenge (Datenmatrix) **D** angewandt wird. Darüber hinaus bezeichnen wir mit O_V die Menge der bezüglich V optimalen Lösungen und konzentrieren uns zunächst auf die Gruppe I_V aller Transformationen T auf $\mathbb{R}$, die die Eigenschaft haben, dass für jede optimale Lösung $L \in O_V$ die transformierte Lösung $T \circ L$ eine optimale Lösung von V bezüglich der transformierten Datenmenge (Datenmatrix) $T \circ$ **D** ist. Auf der anderen Seite müssen wir natürlich die Gruppe T_D aller Transformationen auf $\mathbb{R}$ betrachten, die das Skalenniveau (Messniveau) von **D** repräsentieren. Dann lässt sich die Forderung nach der Adäquatheit von V (bezüglich **D**) in erster Annäherung durch das Postulat $T_D \subset I_V$ präzisieren. Dieses Postulat beinhaltet nur, dass optimale Lösungen bezüglich V auch dann noch optimale Lösungen bezüglich V bleiben, wenn die Daten auf einer zulässigen anderen Skala gemessen werden (z.B. Inches statt Zentimeter, Fahrenheit statt Celsius oder Schweizer Franken statt Euros). Um zu einem Adäquatheitsbegriff zu kommen, der für alle Datenmengen (Datenmatrizen) des gleichen Datentyps gültig ist, muss dann noch von der speziell vorgegebenen

Datenmenge (Datenmatrix) **D** abstrahiert werden. Dazu sei auf Herden und Pallack (2005) verwiesen.

Wir kehren nun zur orthogonalen Form $\text{cov}(X,X) = AA^T + \text{cov}(F,F)$ bzw. $\sum_{XX} = AA^T + E^2$ des Fundamentaltheorems der Faktorenanalyse zurück. Wenn der Korrelationskoeffizient als skalierungsunabhängig gewählt wird, dann gilt für alle bijektiven Abbildungen $B:\mathbb{R} \mapsto \mathbb{R}$ die Gleichung $\text{cov}(B \circ X, B \circ X) = \text{cov}(X,X)$, wobei $B \circ X$ der Zufallsvektor der standardisierten Variablen sei, die aus den beobachteten Variablen nach Verknüpfung mit B hervorgehen. Dies impliziert, dass sämtliche auf dem (orthogonalen) Modell der Faktorenanalyse basierenden Methoden zur Schätzung der Ladungsmatrix A, die auf eine Zerlegung (Faktorisierung) von $\sum_{XX} - E^2$ bzw. E_{XX} hinauslaufen (man denke dabei z.B. an die Hauptkomponentenmethode mit anschließender Varimax-Rotation), invariant sind bezüglich beliebiger bijektiver Abbildungen auf $\mathbb{R}$. All diese Verfahren der Faktorenanalyse sind daher a priori adäquat (bezüglich jeder Datenmenge **D**) und die Frage nach der grundsätzlichen Eignung des zur Schätzung von A herangezogenen faktorenanalytischen Modells tritt zunächst gar nicht auf. Dies scheint ein nicht zu unterschätzender Vorteil skalierungsunabhängiger Korrelationskoeffizienten zu sein.

4.2.4 Über einen Vorschlag zur Behandlung der Kompatibilitätsproblematik der Faktorenanalyse

Begonnen werden soll dieser wichtige Abschnitt mit einigen sehr allgemeinen, nicht zwingend nur für den Korrelationskoeffizienten geltenden, Überlegungen. Seien daher X,Y *endliche* Zufallsvariable, die mit für alle $\omega \in \Omega$ gleicher Wahrscheinlichkeit einen der Werte „0" oder „1" annehmen. Dann ist es plausibel, die *Ähnlichkeit* von X und Y durch eine von der Anzahl der Übereinstimmungen von X und Y abhängenden streng monoton steigenden Abbildung $s:\mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ zu messen. Eine solche Abbildung ist bis auf Ordnungsautomorphismen $T:\left(\left[-1,1\right],\leq\right) \mapsto \left(\left[-1,1\right],\leq\right)$ durch *„1 minus die normierte City-Block-Metrik zwischen X und Y"* gegeben. Diese Beobachtung impliziert sofort, dass eine Abbildung dieses Typs keine Invarianzeigenschaften bezüglich linearer oder gar allgemeiner Transformationen auf $\mathbb{R}$ besitzt. Aus Adäquatheitsgründen (vgl. Teil 2 und 3 dieses Abschnitts) sollen daher hier vor allem Abbildungen $s:\mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ als Ähnlichkeitsmaße zugelassen werden, deren Werte gegenüber möglichst vielen Transformationen auf $\mathbb{R}$ invariant sind. Darüber hinaus sollte „s" selbstverständlich den beim Übergang von der Datenmatrix zur Ähnlichkeitsmatrix (Korrelationsmatrix) entstehenden Informationsverlust möglichst gering halten. Um Letzteres zu erreichen, orientieren wir uns zumindest so weit an obigem Übereinstimmungskriterium, dass die auf $\mathfrak{X}$ gegebene natürliche Präordnung $X \precsim Y \Leftrightarrow P\left(\left\{X \leq Y\right\}\right) = 1$

($X \leq Y$ fast sicher) durch „s" vollständig berücksichtigt werden soll. Für beliebig vorgegebene Zufallsvariablen $X \in \mathfrak{X}, Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ muss „s" also folgende Postulate erfüllen (vgl. hierzu Herden (1987, 1990) und Herden und Pallack (2005, S. 33-34)).

SP1: $X \prec Y \precsim Z \Rightarrow s(X,Z) < s(Y,Z)$

und

SP2: $X \precsim Y \prec Z \Rightarrow s(X,Z) < s(X,Y)$.

Der folgende Satz, dessen Beweis sich im Anhang befindet, unterstreicht, wie stark selbst im allgemeinen Fall die Postulate **SP1** und **SP2** die Möglichkeiten der Konstruktion geeigneter Ähnlichkeitsmaße $s : \mathfrak{X} \times \mathfrak{X} \mapsto [-1,1]$, die plausible Invarianzeigenschaften erfüllen, einschränken. *Es scheint so, dass aus Adäquatheitsgründen wichtige Invarianzeigenschaften für den Preis eines schwer verantwortbaren* ***Informationsverlusts*** *erkauft werden müssen* (vgl. Abschnitt 1).

Satz 4.2:

Es gelten folgende Aussagen:

(i) *Angenommen, für alle Zufallsvariablen $X \in \mathfrak{X}$ und alle affinen Transformationen der Form $m \cdot x + t$ ($m > 0$, t beliebig) auf $\mathbb{R}$ mit $m \cdot X + t \in \mathfrak{X}$ gilt $s(X,X) = s(m \cdot X + t, X)$ oder für alle Zufallsvariablen $X \in \mathfrak{X}$ und alle affinen Transformationen der Form $m \cdot x + t$ ($m > 0$, t beliebig) auf $\mathbb{R}$ mit $m \cdot X + t \in \mathfrak{X}$ gilt $s(X,X) = s(X, m \cdot X + t)$, so kann „s" nur dann wenigstens eines der Postulate* **SP1** *oder* **SP2** *erfüllen, wenn $\mathbb{W}$ höchstens zwei Elemente enthält.*

(ii) *Angenommen, „s" ist symmetrisch und permutationsunabhängig, so kann „s" nur dann wenigstens eines der Postulate* **SP1** *oder* **SP2** *erfüllen, wenn $\mathbb{W}$ höchstens zwei Elemente enthält.*

(iii) *Angenommen, für alle Zufallsvariablen $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ und alle stückweise linearen streng monoton steigenden Transformationen T auf $\mathbb{R}$ mit $T \circ X \in \mathfrak{X}$ und $T \circ Y \in \mathfrak{X}$ gilt $s(X,Y) = s(T \circ X, T \circ Y)$, so kann „s" nur dann wenigstens eines der Postulate* **SP1** *oder* **SP2** *erfüllen, wenn $\mathbb{W}$ höchstens drei Elemente enthält.*

(iv) *Angenommen, für alle Zufallsvariablen $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ und alle affinen Transformationen der Form $m \cdot x + t$ ($m > 0$, t beliebig) auf $\mathbb{R}$ mit $m \cdot X + t \in \mathfrak{X}$ und $m \cdot Y + t \in \mathfrak{X}$ gilt $s(X,Y) = s(m \cdot X + t, m \cdot Y + t)$, so kann „s" nur dann beide Postulate* **SP1** *und* **SP2** *erfüllen, wenn $\mathbb{W}$ höchstens vier Punkte $p_1 < p_2 < p_3 < p_4$ enthält, für die $p_{i+1} - p_i = p_{j+1} - p_j$ für alle $1 \leq i, j \leq 3$ gilt.*

Wir kehren nun zum Spezialfall des Korrelationskoeffizienten zurück, um vollständig die Frage zu klären, unter welchen Bedingungen der Korrelationskoeffizient die Postulate **SP1** und

SP2 erfüllt. Der Vollständigkeit halber ergänzen wir für beliebige Zufallsvariable $X \in \mathfrak{X}, Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ die Postulate **SP1** und **SP2** noch um die schwachen Postulate

WSP1: $X \prec Y \precsim Z \Rightarrow s(X,Z) \leq s(Y,Z)$

und

WSP2: $X \precsim Y \prec Z \Rightarrow s(X,Z) \leq s(X,Y)$.

Darüber hinaus betrachten wir das für Ähnlichkeits- bzw. Distanzmaße oft geforderte *Definitheitspostulat* (man denke z.B. an die Begriffe *Metrik* oder *Norm*).

Def: $s(X,Y) = 1 \Leftrightarrow P\left(\left\{X = Y\right\}\right) = 1$ ($X = Y$ fast sicher)

Um uns dem zu Beginn dieses Abschnitts hervorgehobenen Übereinstimmungskriterium noch mehr anzunähern, wählen wir schließlich noch die Menge $\mathfrak{X}_d$ der dichotomen Zufallsvariablen von $\mathfrak{X}$ aus und bezeichnen mit $\mathfrak{X}_Z$ für jede Zufallsvariable $Z \in \mathfrak{X}$ die Menge derjenigen Zufallsvariablen von $\mathfrak{X}$, die den gleichen Wertebereich wie Z haben. Mit Hilfe dieser Notation lassen sich die folgenden beiden Übereinstimmungspostulate für eine Abbildung $s : \mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ formulieren.

PA1:
$$\forall X \in \mathfrak{X}_d \forall Y \in \mathfrak{X}_d \forall Z \in \mathfrak{X}_d \Big(X \precsim Z \wedge Y \precsim Z \wedge X \in \mathfrak{X}_Z \wedge Y \in \mathfrak{X}_Z \wedge P\left(\left\{X = Z\right\}\right) = P\left(\left\{Y = Z\right\}\right) \Rightarrow s\left(X,Z\right) = s\left(Y,Z\right)\Big)$$

und

PA2:
$$\forall X \in \mathfrak{X}_d \forall Y \in \mathfrak{X}_d \forall Z \in \mathfrak{X}_d \Big(X \precsim Z \wedge Y \precsim Z \wedge X \in \mathfrak{X}_Z \wedge Y \in \mathfrak{X}_Z \wedge P\left(\left\{X = Z\right\}\right) < P\left(\left\{Y = Z\right\}\right) \Rightarrow s\left(X,Z\right) < s\left(Y,Z\right)\Big)$$

Die Postulate **PA1** und **PA2** können wir auch zu folgendem Postulat

PA:
$$\forall X \in \mathfrak{X}_d \forall Y \in \mathfrak{X}_d \forall Z \in \mathfrak{X}_d \Big(X \precsim Z \wedge Y \precsim Z \wedge X \in \mathfrak{X}_Z \wedge Y \in \mathfrak{X}_Z \wedge P\left(\left\{X = Z\right\}\right) \leq P\left(\left\{Y = Z\right\}\right) \Leftrightarrow s\left(X,Z\right) \leq s\left(Y,Z\right)\Big)$$

oder folgendem Postulat

PA*:
$$\forall X \in \mathfrak{X}_d \forall Y \in \mathfrak{X}_d \forall Z \in \mathfrak{X}_d \Big(X \precsim Z \wedge Y \precsim Z \wedge X \in \mathfrak{X}_Z \wedge Y \in \mathfrak{X}_Z \wedge P\left(\left\{X = Z\right\}\right) < P\left(\left\{Y = Z\right\}\right) \Leftrightarrow s\left(X,Z\right) < s\left(Y,Z\right)\Big)$$

zusammenfassen.

Wenn die Postulate **PA1** und **PA2** erfüllt sind, können wir erwarten (die durchgerechneten Beispiele bestätigen dies), dass bei dichotomen Datensätzen eine hohe Übereinstimmung zwischen mit Hilfe von Verfahren der Faktorenanalyse gewonnenen Ergebnissen einerseits und mit Hilfe von Methoden der Clusteranalyse oder der Multidimensionalen Skalierung gewonnenen Ergebnissen andererseits, denen der Euklidische Abstand oder die City-Block-Metrik oder ein verwandtes Distanzmaß zugrunde liegt, bestehen wird. Mit Hilfe der Postulate **PA1** und **PA2** haben wir uns daher auch der Forderung nach Kompatibilität von Methoden der Faktorenanalyse und der Clusteranalyse oder der Multidimensionalen Skalierung angenähert (vgl. Abschnitt 1).

Wir definieren daher: Eine Abbildung $s : \mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ erfüllt das *Kriterium der größtmöglichen Verringerung des beim Übergang von der Datenmatrix zur Ähnlichkeitsmatrix* $\left(s\left(X_i, X_j\right)\right)_{i,j}$ *auftreten könnenden Informationsverlusts,* wenn sie sowohl die *Postulate* **SP1** *und* **SP2** als auch die *Postulate* **PA1** *und* **PA2** erfüllt und darüber hinaus *permutationsunabhängig* ist. Die zusätzliche Forderung der Permutationsunabhängigkeit von „s" haben wir gestellt, um die in Abschnitt 3 beschriebene Adäquatheit von „s" zu erreichen, ohne direkt das strenge Postulat der Skalierungsunabhängigkeit zu fordern.

Der Zusammenhang zwischen den aufgeführten Postulaten liefert folgender Charakterisierungssatz, dessen Beweis sich ebenfalls im Anhang einsehen lässt.

Satz 4.3:

Folgende Aussagen sind äquivalent:

(i) *Der Korrelationskoeffizient erfüllt* **WSP1**.

(ii) *Der Korrelationskoeffizient erfüllt* **WSP2**.

(iii) *Der Korrelationskoeffizient erfüllt* **SP1**.

(iv) *Der Korrelationskoeffizient erfüllt* **SP2**.

(v) *Der Korrelationskoeffizient erfüllt sowohl* **PA1** *als auch* **PA2** *und* $\mathbb{W}$ *enthält höchstens zwei Elemente.*

(vi) *Der Korrelationskoeffizient erfüllt* **Def**.

(vii) $\mathbb{W}$ *enthält höchstens zwei Elemente.*

Im Zusammenhang mit Satz 4.1 besagt Satz 4.3, dass ein genau zwei Elemente (Werte) umfassender Wertebereich $\mathbb{W}$ *eine größtmögliche Verringerung des beim Übergang von der Datenmatrix zur Korrelationsmatrix auftreten könnenden Informationsverlusts garantiert.*

Die grundsätzliche Bedeutung der Übereinstimmungspostulate **PA1** und **PA2** ist über Satz 4.3 hinaus vor allen Dingen darin begründet, dass mit ihrer Hilfe folgender Charakterisierungssatz hergeleitet werden kann. Auch dieser Beweis findet sich im Anhang.

Satz 4.4:
Sei $s : \mathfrak{X} \times \mathfrak{X} \mapsto [-1,1]$ *eine symmetrische Abbildung. Dann sind folgende Aussagen äquivalent.*

(i) *„s" erfüllt das Kriterium der größtmöglichen Verringerung des beim Übergang von der Datenmatrix zur Ähnlichkeitsmatrix* $\left(s\left(X_i, X_j\right)\right)_{i,j}$ *auftreten könnenden Informationsverlusts.*

(ii) *Für* $\mathbb{W}$ *und „s" gelten folgende Bedingungen:*

W: $\mathbb{W}$ *enthält höchstens zwei Elemente* u *und* v.

B1: $s(X,Y) = s(-X+u+v, -Y+u+v)$ *für alle Zufallsvariablen* $X \in \mathfrak{X}$ *und* $Y \in \mathfrak{X}$.

B2: *Zu allen Zufallsvariablen* $Z \in \mathfrak{X}$ *existiert eine ordnungserhaltende Abbildung* $T_Z : ([-1,1], \leq) \mapsto ([-1,1], \leq)$, *so dass für alle Zufallsvariablen* $X \in \mathfrak{X}$ *mit* $X \precsim Z$ *die Gleichung* $s(X,Z) = T_Z\left(cor(X,Z)\right)$ *erfüllt ist.*

Fassen wir obige Sätze zusammen, wobei wir insbesondere an Satz 4.3 und Satz 4.4 denken, so können wir festhalten, dass sich eine (in unserem Sinne) größtmögliche Verringerung des beim Übergang von der Datenmatrix zur Korrelationsmatrix (möglicherweise) entstehenden Informationsverlusts durch Dichotomisierung der beobachteten Daten erreichen lässt. Durch das Umrechnen der Ausgangsdaten in „0-1-Daten" kann also, wie obige Sätze unterstreichen, die Kompatibilitätsproblematik der Faktorenanalyse weitestgehend vermieden werden. Man vergleiche ergänzend hierzu auch die dichotome Daten studierende Arbeit von GLÖCKNER-RIST und HOIJTINK (2003).

Aus diesen Sätzen resultiert daher folgender Vorschlag für eine konkrete Vorgehensweise, der anhand der zu Anfang gewählten Bewertungsskala –3, –2, –1, 0, 1, 2, 3 beschrieben sei. Im ersten Schritt übersetzen wir dazu die Skala –3, –2, –1, 0, 1, 2, 3 im Sinne der von WILLE (1982) entwickelten formalen Begriffsanalyse in einen einwertigen Kontext (vgl. hierzu auch GANTER und WILLE (1996) und die hier im Zusammenhang mit mehrwertigen Kontexten entwickelte *begriffsanalytische Messtheorie*). Die Gegenstände dieses Kontexts sind in den von uns betrachteten Fällen die beobachteten Zufallsvariablen, während seine Merkmale die Probanden sind, denen Tests oder Bewertungsskalen vorgelegt wurden. Der auf diese Weise vorgegebene formale Kontext wird dadurch mehrwertig, dass die Probanden verschiedene Punktzahlen erreichen oder verschiedene Bewertungen vornehmen können. Betrachten wir z.B. die dem ersten Beispiel zugrunde liegende Bewertungsskala –3, –2, –1, 0, 1, 2, 3, so entsprechen jedem Probanden 7 verschiedene *Merkmalsausprägungen*, von denen sie oder er sich für eine entscheidet. Im Sinne der begriffsanalytischen Messtheorie lässt sich die Bewertungsskala –3, –2, –1, 0, 1, 2, 3 daher in folgenden *einwertigen Kontext* übersetzen.

≤	-3	-2	-1	0	1	2	3
-3	1	1	1	1	1	1	1
-2	0	1	1	1	1	1	1
-1	0	0	1	1	1	1	1
0	0	0	0	1	1	1	1
1	0	0	0	0	1	1	1
2	0	0	0	0	0	1	1
3	0	0	0	0	0	0	1

Abbildung 4.9 Einwertiger Kontext der Bewertungsskala

-3 wird also durch das Tupel (1,1,1,1,1,1,1), -2 durch das Tupel (0,1,1,1,1,1,1), -1 durch das Tupel (0,0,1,1,1,1,1) … und schließlich 3 durch das Tupel (0,0,0,0,0,0,1) dichotomisiert. Selbstverständlich hätten wir wegen Satz 4.1 statt dieses, die Relation ≤ wiedergebenden, Kontexts auch den dualen, die Relation ≥ wiedergebenden, Kontext wählen können. Dies bedeutet insbesondere, dass die in Abschnitt 1 beschriebenen Äquivalenzklassen [**D**] stets aus genau zwei Datenmatrizen bestehen, die durch Vertauschung von „0" und „1" (allgemein: des kleineren und des größeren Werts) auseinander hervorgehen (vgl. die entsprechenden Erörterungen im Anschluss an die Vorstellung der Bedingungen **Z1** und **Z2**).

Alleine diese einfache Beobachtung verdeutlicht (man vergleiche die Zahl 2 mit der u. U. gigantischen Zahl $2 \cdot \binom{v}{2}^n$ *), dass die Dichotomisierung der Variablen mit Hilfe begriffsanalytischer Messtheorie den Informationsverlust beim Übergang von der Datenmatrix der beobachteten Variablen zur entsprechenden Korrelationsmatrix praktisch verschwinden lässt.*

Die Datenmatrix **D1** übersetzt sich obigem Kontext entsprechend in folgende 8×57 Datenmatrix **D3**:

	***P*₁**							…	***P*₈**						
X_1	0	0	0	0	0	0	1	…	0	0	0	0	0	0	1
X_2	0	0	0	0	0	1	1	…	0	0	0	0	0	0	1
X_3	1	1	1	1	1	1	1	…	0	1	1	1	1	1	1
X_4	0	0	0	0	0	1	1	…	0	0	0	0	0	0	1
X_5	1	1	1	1	1	1	1	…	0	0	1	1	1	1	1
X_6	0	0	0	0	0	1	1	…	0	0	0	0	0	0	1
X_7	1	1	1	1	1	1	1	…	1	1	1	1	1	1	1

Abbildung 4.10 Datenmatrix D3

Die Möglichkeit der automatischen Umrechnung von in Excel erstellten Datenmatrizen, die auf Bewertungsskalen oder Punktbewertungen basieren, in „0-1-Datenmatrizen" wurde speziell für dieses Projekt entwickelt, steht zur Verfügung und wird auch im weiteren Verlauf der Arbeit mehrfach seine Anwendung finden.

Betrachten wir z.B. die Variable X_4 und X_7, so errechnet sich ihr Korrelationskoeffizient als

$$cor\left(X_4, X_7\right) =$$

cor ((0, 0, 0, 0, 0,1,1, 0, 0, 0, 0, 0,1,1, 0, 0, 0, 0, 0, 0,1, 0, 0, 0, 0, 0,1,1, 0, 0, 0, 0, 0,1,1, 0, 0, 0, 0, 0,1,1, 0, 0, 0, 0, 0, 0,1, 0, 0, 0, 0, 0, 0,1), (0, 0,1,1,1,1,1, 0, 0,1,1,1,1,1, 0, 0, 0, 0,1,1,1, 0, 0, 0, 0,1,1,1, 0, 0,1,1,1,1,1, 0, 0,1,1,1,1,1, 0, 0, 0, 0,1,1,1, 0, 0,1,1,1,1,1)).

Auf die so errechnete und unten aufgeführte Korrelationsmatrix **Corr2** lassen sich dann beliebige Verfahren der Faktorenanalyse zur Schätzung der Ladungsmatrix A anwenden.

cor	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1						
X_2	0,887	1					
X_3	0,118	0,118	1				
X_4	0,899	0,899	0,131	1			
X_5	0,231	0,231	0,510	0,256	1		
X_6	0,899	0,899	0,131	0,899	0,256	1	
X_7	0,118	0,118	0,648	0,131	0,510	0,131	1

Abbildung 4.11 Korrelationsmatrix der Datenmatrix D3

Der Vergleich von **Corr1** mit **Corr2** verdeutlicht die augenfällige Verringerung des Informationsverlusts, der beim Übergang von der Datenmatrix zur Korrelationsmatrix durch Dichotomisierung der Ausgangsdaten gelungen ist. Insbesondere weist **Corr2** auf die vermutete Existenz der durch die Zufallsvariablen X_1, X_2, X_4 und X_6 bzw. X_3, X_5 und X_7 repräsentierten Faktoren hin.

Ergänzend sei an dieser Stelle noch die die durch Dichotomisierung der Datenmatrix **D1*** gewonnene Korrelationsmatrix **Corr2*** aufgeführt.

cor	A_1	L_1	A_2	L_2	A_3	L_3
A_1	1					
L_1	0,662	1				
A_2	0,942	0,703	1			
L_2	0,662	1	0,703	1		
A_3	0,948	0,734	0,893	0,734	1	
L_3	0,734	0,948	0,779	0,948	0,80	1

Abbildung 4.12 Korrelationsmatrix der dichotomisierten Datenmatrix D1*

Auch hier ist erkennbar, dass nun die Ähnlichkeiten der Ausgangsdaten als hoch angesehen werden, also im Wesentlichen korrekt interpretiert werden.

Zusammengefasst wird deutlich, wie eine sofortige praktische Umsetzung des gemachten Vorschlags möglich ist:

(1) Man erfasst die auf einer vorgegebenen Bewertungsskala (Ratingskala) oder allgemeiner Punkteskala gemessenen Bewertungen oder Punkte wie gewöhnlich, also unverändert, in einer Excel-Datei.

(2) Man benutzt das erstellte Programm, um die in (1) gewonnenen Ausgangsdaten in „0-1-Daten" zu transformieren.

(3) Man verwendet die in (2) ermittelte „0-1-Datenmatrix", um mit Hilfe von SPSS ein *geeignetes* faktorenanalytisches Verfahren zur Auswertung der Daten anzuwenden. *Da beschriebene Überlegungen von der Hauptkomponentenmethode mit anschließender Varimax-Rotation ausgingen, empfiehlt sich zunächst die Benutzung dieses Verfahrens.*

4.3 Die Verteilung der Probanden auf die latenten Faktoren

Basierend auf der eben vorgestellten Variante der Faktorenanalyse gelingt es nun die Dimensionen eines auf Rangdaten basierenden Datensatzes ohne spürbaren Informationsverlust zu reduzieren und latente Faktoren zu bestimmen. Sind nun derartige, den Datensatz beschreibende, Faktoren extrahiert worden, gilt es die Probanden entsprechend ihres Abstimmungsverhaltens auf die identifizierten Faktoren zu verteilen. Es soll also konkret die Frage beantwortet werden, welcher Proband welchen Faktor favorisiert.

Wir betrachten dazu ein erläuterndes Beispiel, welches dem ersten Beispiel aus dem Abschnitt über die mathematischen Grundlagen der Faktorenanalyse nachempfunden ist.

Heute wirst du zum Fernsehtester und darfst die folgenden TV-Sendungen bewerten, indem du die Sendungen danach ordnest, wie gerne du sie schaust.

DEUTSCHLAND SUCHT DEN SUPERSTAR
DAS SUPERTALENT
WISSEN MACHT AH!
UNSER STAR FÜR BAKU
GALILEO
THE VOICE OF GERMANY
GALILEO MYSTERY

Du darfst auch verschiedenen Sendungen die gleiche Platzierung geben.
Deine Lieblingssendung bekommt den ersten Platz!

Abbildung 4.13 Fragebogen für das erläuternde Beispiel

Im Folgenden gehen wir davon aus, dass acht zufällig ausgewählte Probanden die in Abbildung 4.13 formulierte Aufgabe umsetzen, indem sie die sieben Fernsehsendungen jeweils in eine individuelle Prioritätsordnung überführen, wobei die Zahlen „8” und „7” auch in diesem Beispiel zufällig gewählt und ohne Bedeutung für die weiteren Ausführungen sind.

Die so gewählte Rangordnung ist lückenfrei, lässt aber verbundene Ränge zu, was in diesem Fall bedeutet, dass mehrere Fernsehsendungen den gleichen Rang bekommen können. Wird eine Platzierung, zum Beispiel der zweite Platz, mehrfach besetzt, soll die Rangfolge, anders als beispielsweise bei Sporttabellen üblich, dennoch mit dem nächstgrößeren, hier dem Dritten, Rang fortgesetzt werden.

Lückenfreiheit ist immer dann angezeigt, wenn die Reihenfolge lediglich ordinale Signifikanz hat, was bedeutet, dass lediglich die Rangfolge der Fernsehsendungen, nicht aber irgendwie quantifizierbare Qualitätsunterschiede zwischen den TV-Sendungen, beobachtet werden können. Da die Probanden dazu aufgefordert wurden eine Rangfolge zu erstellen, was bedeutet, dass jede subjektive Bewertung der einzelnen Probanden lediglich ordinale Signifikanz hat, darf auch die resultierende Rangordnung nur ordinale Signifikanz haben.

Bei ähnlichen Fragestellungen werden daher auch im Verlauf der empirischen Studie dieser Arbeit lückenfreie Rangordnungen verwendet, bei denen verbundene Ränge zugelassen sind. Die Argumentation verläuft stets analog.

Im Folgenden werden die obige Einflussfaktoren repräsentierenden Zufallsvariablen nacheinander mit X_1 (Deutschland sucht den Superstar), X_2 (Das Supertalent), X_3 (Wissen macht Ah!), X_4 (Unser Star für Baku), X_5 (Galileo), X_6 (The Voice of Germany) und X_7 (Galileo Mystery) abgekürzt. Wir stellen uns vor, dass das Ergebnis dieser Befragung in nachfolgender Datenmatrix **D1** erfasst ist.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	1	1	1	1	4	3	6	2
X_2	1	2	1	2	3	4	4	3
X_3	2	5	1	3	2	5	2	2
X_4	1	3	1	3	3	7	3	2
X_5	2	6	2	3	2	2	2	1
X_6	1	4	1	3	2	6	1	2
X_7	2	7	2	4	1	1	5	1

Abbildung 4.14 Datenmatrix D1

Bevor bestimmt werden kann, welcher Proband welches Fernsehprogramm am liebsten schaut, müssen in dem vorliegenden Beispiel zunächst, basierend auf der Datenmatrix **D1**, die dem Fernsehverhalten zugrunde liegenden latenten Faktoren überhaupt erst extrahiert werden.

Nach Transformation von **D1** in die zugehörige 0-1-Datenmatrix kann die Faktorenanalyse (hier Hauptkomponentenmethode mit anschließender Varimax-Rotation) wie erläutert angewendet werden. Dabei ergibt sich, wie erwartet, folgendes Ergebnis:

Rotierte Komponentenmatrix^a

	Komponente	
	1	2
Supertalent	,864	
Unser Star für Baku	,817	
DSDS	,750	
The Voice	,629	
Galileo		,913
Galileo Mystery		,782
Wissen macht Ah!		,733

Abbildung 4.15 Rotierte Komponentenmatrix

Offensichtlich erhalten wir in diesem Beispiel zwei zu extrahierende Faktoren. Der erste Faktor beinhaltet die verschiedenen Casting-Sendungen, während der zweite Faktor diejenigen

Sendungen enthält, die sich um Bildung und Wissensvermittlung bemühen. Dementsprechend bezeichnen wir den ersten Faktor als *Casting* und den zweiten Faktor als *Bildung*.

Nun kann untersucht werden, welche der acht Probanden lieber Sendungen des Faktors *Casting* schauen und welche der acht Probanden eher Sendungen des Faktors *Bildung* favorisieren. Dies soll mit Hilfe clusteranalytischer Verfahren geschehen.

Dazu wird der Datensatz im ersten Schritt entsprechend der gewonnenen Faktoren sortiert, so dass sich folgende Datenmatrix **D2** ergibt:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	1	1	1	1	4	3	6	2
X_2	1	2	1	2	3	4	4	3
X_4	1	3	1	3	3	7	3	2
X_6	1	4	1	3	2	6	1	2
X_3	2	5	1	3	2	5	2	2
X_5	2	6	2	3	2	2	2	1
X_7	2	7	2	4	1	1	5	1

Abbildung 4.16 Datenmatrix D2

Die blau markierten Zeilen geben die Ränge der Sendungen des Faktors *Casting* an, die orange markierten Zeilen stehen für die Ränge der Sendungen des Faktors *Bildung*.

Bei Betrachtung der Datenmatrix **D2** fällt auf, dass die Probanden P_1, P_2, P_3 und P_4 die Sendungen des ersten Faktors besonders gut bewertet haben. Die *Bildungs*-Sendungen werden durchgehend schlechter beurteilt. Diese Beobachtung legt die Vermutung nahe, dass diese vier Probanden lieber *Casting*-Sendungen schauen, als sich vom Fernsehen bilden zu lassen. Im Gegensatz dazu haben die Probanden P_5, P_6, P_7 und P_8 die *Bildungs*-Sendungen insgesamt besser bewertet als die übrigen. Hier lässt sich dementsprechend vermuten, dass diese Probanden lieber *Bildungsfernsehen* konsumieren.

Diese Vermutung gilt es, durch eine Clusteranalyse zu bestätigen. Die Clusteranalyse ist aber eigentlich

> *„kein einzelnes Verfahren, sondern ein Oberbegriff für eine Gruppe von statistisch ausgesprochen heterogenen Verfahren, die es ermöglichen, Objekte anhand ihrer Merkmalsausprägungen zu Gruppen ('Clustern') so zusammenzufassen, dass einerseits die Ähnlichkeit zwischen den Objekten innerhalb dieser Cluster möglichst groß ist (Ziel 1: hohe Intraclus-*

> *ter-Homogenität), aber anderseits die Ähnlichkeit zwischen den Clustern möglichst gering ist (Ziel 2: geringe Intercluster-Homogenität)"* (Schendera 2010, S. 8).

Dementsprechend ist es zunächst entscheidend ein geeignetes clusteranalytisches Verfahren auszuwählen, welches zu den vorhandenen Daten passt. Dabei sind vor allem zwei Aspekte zu beachten:

> (i) *Die Wahl eines Fusionierungsalgorithmus*: Zunächst gilt es, aus einer Reihe von Fusionierungsalgorithmen ein für die erhobenen Daten passendes Verfahren auszuwählen. Dabei wird grundsätzlich zwischen hierarchischen Clusteranalysen, bei denen eine sukzessive Aufteilung in Cluster vorgenommen wird, und Clusterzentrenanalysen, bei denen die Anzahl der Gruppen vorab vorgegeben wird, unterschieden. Verfahren der hierarchischen Clusteranalyse sind unter anderem das Single-Linkage Verfahren, das Complete-Linkage Verfahren oder der Ward-Algorithmus. Das Single-Linkage Verfahren ermittelt die Cluster über den geringsten Abstand zweier Elemente, neigt aber wegen schwacher Homogenität zur Kettenbildung und tendiert dazu, Ausläufer von lang gezogenen oder unregelmäßigen Verteilungen abzuschneiden. Das Complete-Linkage Verfahren ermittelt die Cluster über den Maximalabstand aller Elemente, was wiederum dazu führt, dass aufgrund von strenger Homogenität viele kleine Gruppen gebildet werden. Der Ward-Algorithmus generiert die Cluster letztlich über den minimalen Anstieg der Intraclustervarianz und liefert damit überlappungsfreie Clusterstrukturen (vgl. ebd., S. 25 f).

> (ii) *Die Wahl des Proximitätsmaßes*: Ist ein geeigneter Fusionierungsalgorithmus ausgewählt worden, muss ein dazu adäquates Proximitätsmaß bestimmt werden. Das Proximitätsmaß dient dazu, die Ähnlichkeit zweier Objekte bzw. Fälle im Bezug auf die clusterbildenden Eigenschaften durch einen Zahlenwert auszudrücken, wobei Ähnlichkeits- und Distanzmaße unterschieden werden können. Bei einem Ähnlichkeitsmaß drückt ein hoher Wert auch eine hohe Ähnlichkeit aus, während bei einem Distanzmaß nach den kürzesten Abständen gesucht wird. Bei der Wahl des Proximitätsmaßes muss vor allem das Skalenniveau beachtet werden (vgl. Tauberger 2008, S. 283). Besonders bekannte Distanzmaße sind zum Beispiel der (quadrierte) Euklidische Abstand, die City-Block-Metrik und die Minkowski-Distanz. Aus der Fülle an Ähnlichkeitsmaßen sticht im Bezug auf Verfahren der Clusteranalyse vor allem die Pearson-Korrelation heraus. Es gilt zu beachten, dass verschiedene Proximitätsmaße zu unterschiedlichen Ergebnissen führen können.

Da dementsprechend auch Methoden der Clusteranalyse auf einem Distanzmaß basieren, müssen die erhobenen Daten auch hier entsprechend des Forschungsansatzes dieser Arbeit in

0-1-Daten übertragen werden. Dazu wird für jeden Probanden P_i $\left(1 \le i \le 8\right)$ ein Vektor der folgenden Bauart generiert:

$$P_i = \left(\underbrace{\underbrace{1,1,1,\ldots,1,0,0,0,\ldots,0}_{\text{geordnete 0-1-Daten von Faktor 1}}, \underbrace{1,1,1,\ldots,1,0,0,0,\ldots,0}_{\text{geordnete 0-1-Daten von Faktor 2}}}_{\text{sämtliche 0-1-Daten des Probanden } P_i} \right)$$

Zur Erstellung dieses obigen Vektors werden die 0-1-Daten für die Sendungen der beiden Faktoren zunächst getrennt voneinander betrachtet. Sämtliche 0-1-Daten des ersten Faktors werden aneinandergereiht und so sortiert, dass alle Einsen vorne und alle Nullen hinten stehen. Das Vorgehen für den zweiten Faktor verläuft analog. Im Anschluss daran werden beide Datenreihen zu einer gemeinsamen zusammengefasst. Durch diese Maßnahme wird jeweils die grundsätzliche Bewertung der beiden Faktoren durch die Probanden dargestellt. Dies meint, dass durch die unterschiedlichen Anzahlen der Einsen deutlich werden sollte, welchen Faktor der jeweilige Proband bevorzugt.

Auf diese acht Datenreihen lässt sich nun das Complete-Linkage Verfahren anwenden, da dieses invariant ist gegenüber ordinalen Transformationen. Bei dieser Methode werden diejenigen Probanden zu einem Cluster zusammengefasst, die eine höhere Ähnlichkeit aufweisen, als zu den Probanden der anderen Cluster. Als Distanzmaß bietet sich der Euklidische Abstand an, weil dieser bei 0-1-Bewertungen die Anzahl der Nichtübereinstimmungen misst. Da wir ordinales Abstandsniveau voraussetzen, ist dieses Distanzmaß lediglich bis auf ordinale Transformationen eindeutig bestimmt.

Nun liefert das Verfahren das in Abbildung 4.17 dargestellte Ergebnis.

Cluster-Zugehörigkeit

Fall	2 Cluster
P1	1
P2	2
P3	1
P4	1
P5	1
P6	1
P7	1
P8	1

Abbildung 4.17 Cluster-Zugehörigkeit (1)

Dabei werden die Probanden $P_1, P_3, P_4, P_5, P_6, P_7$ und P_8 dem ersten Cluster zugeordnet, einzig P_2 ist Inhalt des zweiten Clusters. Bis auf den Probanden P_2 schauen demnach sämtliche Probanden lieber Casting-Sendungen. Dieses Ergebnis widerspricht allerdings der Vermutung und entspricht damit nicht dem tatsächlichen Abstimmungsverhalten der Probanden.

Die Erklärung für die *falschen* Cluster-Zugehörigkeiten ist die Folgende: Betrachten wir noch einmal die Datenmatrix **D2**, dann fällt auf, dass einige der Probanden (zum Beispiel P_1, P_3 und P_8) insgesamt ein äußerst positives Abstimmungsverhalten gezeigt und sämtliche zur Auswahl stehenden Sendungen aufgrund vieler verbundener Ränge eher gut bewertet haben, während die übrigen Probanden (zum Beispiel P_2 und P_6) die Bewertungsskala nahezu vollständig ausschöpften. Dieses Ereignis führt dazu, dass die durch die Clusteranalyse gebildeten Probandengruppen danach zusammengesetzt werden, ob sie eine eher positive oder eine eher ausgewogene Gesamtbewertung aller Sendungen abgegeben haben. Das Ziel ist es aber, festzustellen, ob die Sendungen des ersten Faktors besser oder schlechter bewertet wurden als die Sendungen des zweiten Faktors.

Um dieses Ziel zu erreichen, müssen die äußerst positiven Bewertungen denjenigen Bewertungen angeglichen werden, die die Skala weitestgehend vollständig ausschöpfen.

Die Nichtübereinstimmung kommt augenscheinlich durch die in dem Abstimmungsverhalten beobachteten Lücken, die einem ordinalen Skalenniveau ohnehin widersprechen, zustande, so dass eine dem Datenniveau angepasste Analyse die Lücken schließt. Eine solche Anpassung lässt sich erreichen, indem wir uns bei verbundenen Rängen von einer lückenfreien Rangordnung verabschieden. Wird also der zweite Platz doppelt belegt, wird die Rangfolge mit dem vierten Platz fortgeführt.

Dieses Vorgehen liefert die Datenmatrix **D3**:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
X_1	1	1	1	1	7	3	7	3
X_2	1	2	1	2	5	4	5	7
X_4	1	3	1	3	5	7	4	3
X_6	1	4	1	3	2	6	1	3
X_3	5	5	1	3	2	5	2	3
X_5	5	6	6	3	2	2	2	1
X_7	5	7	6	7	1	1	6	1

Abbildung 4.18 Datenmatrix D3

Betrachtet man **D3,** wird deutlich, dass unsere obige Vermutung, nämlich dass die Probanden P_1, P_2, P_3 und P_4 zu dem Casting-Show favorisierenden Cluster und die Probanden P_5, P_6, P_7 und P_8 zu dem bildungsnahen Cluster gehören, weiterhin Gültigkeit besitzt. Im Unterschied zu **D2** wird nun allerdings von allen Probanden die Bewertungsskala weitestgehend ausgereizt.

Das weitere Vorgehen entspricht nun wieder der oben beschriebenen Idee, für jeden Probanden die 0-1-Daten zunächst nach Faktorenzugehörigkeit und im Anschluss nach Einsen und Nullen zu sortieren. Zur Bestimmung der Cluster-Zugehörigkeit der Probanden wird erneut das Complete-Linkage Verfahren auf Basis des quadrierten Euklidischen Abstands verwendet.

Wir erhalten nun folgendes Ergebnis:

Cluster-Zugehörigkeit

Fall	2 Cluster
P1	1
P2	1
P3	1
P4	1
P5	2
P6	2
P7	2
P8	2

Abbildung 4.19 Cluster-Zugehörigkeit (2)

Wie bereits vermutet und nun durch ein Verfahren der Clusteranalyse bestätigt, tendieren die ersten vier Probanden zu Faktor 1, während die letzten vier Probanden den zweiten Faktor bevorzugen.

Die verwendete Vergröberung der Skala bei verbundenen Rängen führt dementsprechend zu einem zufriedenstellenden Ergebnis, so dass zusammengefasst deutlich wird, wie eine sofortige praktische Umsetzung des gemachten Vorschlags möglich ist:

(1) Man bestimmt die latenten Faktoren eines Datensatzes mit Hilfe der 0-1-Datenmatrix.

(2) Man ordnet den Datensatz entsprechend der extrahierten Faktoren.

(3) Man vergröbert die Skala bei denjenigen Probanden, die verbundene Ränge genutzt haben.

(4) Man sortiert für jeden Probanden die 0-1-Daten zunächst nach Faktorenzugehörigkeit und im Anschluss nach Einsen und Nullen.

(5) Zur Bestimmung der Cluster-Zugehörigkeit der Probanden wird das Complete-Linkage Verfahren auf Basis des Euklidischen Abstands verwendet.

4.4 Erweiterung des Hauptsatzes der formalen Begriffsanalyse

Die begriffsanalytische Skalierung im Sinne der formalen Begriffsanalyse nach Wille (1982) und Ganter und Wille (1996) ist bereits im Zusammenhang mit den Korrelationsmatrizen in Abschnitt 4.2 dieser Arbeit aufgetreten. Dabei ist jedes einzelne Merkmal skaliert. Werden diese Skalen in eine entsprechende Kreuztabelle umgearbeitet, so wird aus dem mehrwertigen Kontext ein einwertiger Kontext.

Um nun aber zu überprüfen, inwieweit das Ergebnis der auf den 0-1-Daten basierenden Faktorenanalyse methodenunabhängig ist, soll eine modifizierte Version der formalen Begriffsanalyse ebenfalls auf den Datensatz angewendet werden, weil die formale Begriffsanalyse die Erzeugung und Visualisierung der Begriffshierarchien auf einer mathematisch fundierten Basis bietet und einen durch die Faktorenanalyse möglichen Informationsverlust a priori ausschließt, da eben keine Ähnlichkeitsmatrix dem Verfahren zugrunde liegt.

In vorliegendem Fall soll die formale Begriffsanalyse hierbei auf beliebige transitive, antisymmetrische und reflexive Strukturen erweitert werden, indem wir nicht nur vollständigen Verbänden entsprechende Begriffsverbände zuordnen, sondern jeder geordneten Menge eine entsprechende Unterstruktur der Begriffsverbände zuweisen, die aus *einfach erzeugten* Begriffen besteht. Sei dazu $\mathcal{K}\left(G,M,I\right)$ ein formaler Kontext. Dann heißt ein Begriff $\left(C,D\right)$ des zugehörigen Begriffsverbunds $\left(\mathcal{B}\left(G,M,I\right),\preceq\right)$ *einfach erzeugt*, wenn ein Gegenstand $g \in C$ existiert, so dass $D=\left\{g\right\}^{\perp}=\left\{m \in M \mid \left(g,m\right)\in I\right\}$ und $C=\left\{g\right\}^{\perp\perp}=\left\{g' \in G \mid \left(g',m\right)\in I \text{ für alle } m \in D\right\}$ gilt. Die Unterstruktur der einfach erzeugten Begriffe von $\left(\mathcal{B}\left(G,M,I\right),\preceq\right)$ sei im Folgenden mit $\left(\mathcal{E}\left(G,M,I\right),\preceq\right)$ abgekürzt.

Erweiterung des Hauptsatzes der formalen Begriffsanalyse auf beliebige Ordnungen

Sei X eine nicht-leere Menge und R eine binäre Relation auf X. Dann wiederholen wir zunächst, dass R genau dann eine (teilweise) Ordnung auf X ist, wenn R reflexiv, antisymmetrisch und transitiv ist.

Auf Basis dieser Begriffsbildung lässt sich der Hauptsatz der formalen Begriffsanalyse durch folgenden Satz, dessen Beweis sich im Anhang einsehen lässt, ergänzen.

Satz 4.5:

Es gelten folgende Aussagen:

(i) *Sei $\mathcal{K}\left(G,M,I\right)$ ein formaler Kontext. Dann ist $\left(\mathcal{E}\left(G,M,I\right),\preceq\right)$ eine (teilweise) geordnete Menge.*

(ii) *Sei X eine nicht-leere Menge und R eine binäre Relation auf X. Dann ist R genau dann eine (teilweise) Ordnung auf X, wenn ein formaler Kontext $\mathcal{K}\left(G,M,I\right)$ existiert, so dass $\left(X,R\right)$ zu $\left(\mathcal{E}\left(G,M,I\right),\preceq\right)$ ordnungsisomorph ist.*

Vorgehen in praktischen Anwendungsfällen:

Die praktische Umsetzung dieses Satzes soll im Folgenden anhand eines kleinen Beispiels erläutert werden. Wir nehmen dazu an, dass sechs Probanden (P_1 bis P_6) vier verschiedene Gegenstände (X_1 bis X_4) entsprechend ihrer Beliebtheit in eine Rangfolge gebracht haben. Dabei entspricht der erste Platz dem beliebtesten und der vierte Platz dem unbeliebtesten Gegenstand. Verbundene Ränge sind zugelassen. Dabei sei folgender Datensatz entstanden:

	P_1	**P_2**	**P_3**	**P_4**	**P_5**	**P_6**
X_1	1	1	1	1	4	1
X_2	1	2	1	2	3	2
X_3	2	4	1	3	2	3
X_4	1	3	1	3	1	4

Abbildung 4.20 Beispieldatensatz formale Begriffsanalyse

Bereits auf den ersten Blick fällt auf, dass X_1 besonders gut bewertet wurde. Fünf der sechs Probanden weisen diesem Gegenstand die Spitzenposition zu. Während X_2 ebenfalls konstant gut bewertet wird, gilt dies vor allem für die letzten beiden Gegenstände eher im negativen Sinne. Das Verfahren der formalen Begriffsanalyse sollte dementsprechend diese hierarische Ordnung zum Ausdruck bringen und ähnlich bewertete Gegenstände identifizieren.

Das erweiterte begriffsanalytische Verfahren zur Analyse des Datensatzes verläuft nun in vier Schritten. Der erste Schritt verfolgt das Ziel, die Nachfolgermengen der Gegenstände basierend auf den Probandenbewertungen zu bestimmen. Dazu zählen wir zunächst für alle $i\in\left\{1,\ldots,4\right\}$ und alle $j\in\left\{1,\ldots,4\right\}$, wie oft Gegenstand X_i besser oder gleich gut wie Gegenstand X_j $\left(i\neq j\right)$ bewertet wurde und halten das Ergebnis in nachfolgender Tabelle fest:

$\succeq$	**X_1**	**X_2**	**X_3**	**X_4**
X_1		**5**	5	5
X_2	3		5	5
X_3	**2**	**2**		3
X_4	3	3	5	

Abbildung 4.21 Nachfolgermengen bestimmen (1)

In dem betrachteten Fall haben beispielsweise fünf Probanden X_1 besser als oder gleich gut wie X_2 einsortiert (fett markiert), aber nur zwei Probanden X_3 für besser oder gleich gut wie X_1 und X_2 befunden (fett markiert).

Wenn mehr Probanden den Gegenstand X_i mindestens so gut bewertet haben wie den Gegenstand X_j als umgekehrt setzen wir an die entsprechende Stelle in der Tabelle ein Kreuz, welches die Dominanz des einen über den anderen Gegenstand symbolisiert.

$\succeq$	**X_1**	**X_2**	**X_3**	**X_4**
X_1		X	X	X
X_2			X	X
X_3				
X_4			X	

Abbildung 4.22 Nachfolgermengen bestimmen (2)

In diesem Beispiel halten fünf Probanden X_1 für besser oder gleich gut wie X_2, wobei X_2 lediglich bei drei Probanden eine höhere oder vergleichbare Beliebtheit wie X_1 besitzt.

Erst jetzt können die Nachfolgermengen $N(X_i) := \left\{ X_j \mid (X_i, X_j) \text{ besitzt Kreuz} \right\}$ endgültig bestimmt werden. Diese beinhalten logischerweise diejenigen Gegenständen, die sich einer geringen Beliebtheit erfreuen:

$N(X_1)$	**$N(X_2)$**	**$N(X_3)$**	**$N(X_4)$**
X_2	X_3		X_3
X_3	X_4		
X_4			

Abbildung 4.23 Nachfolgermengen bestimmen (3)

An dieser Stelle beginnt der zweite Schritt, an dessen Ende die Nachfolgermengen in einer Kreuztabelle dargestellt werden sollen. Dazu wird lediglich überprüft, welche Nachfolgermengen Teilmenge einer anderen sind. Beispielsweise sind alle Gegenstände der Mengen $N(X_2)$, $N(X_3)$ und $N(X_4)$ auch in der Menge $N(X_1)$ enthalten, so dass diese Nachfolgermengen allesamt Teilmengen von $N(X_1)$ sind. Die entsprechende Kreuztabelle stellt sich wie folgt dar:

$\subset$	**$N(X_1)$**	**$N(X_2)$**	**$N(X_3)$**	**$N(X_4)$**
$N(X_1)$				
$N(X_2)$	X			
$N(X_3)$	X	X		X
$N(X_4)$	X	X		

Abbildung 4.24 Kreuztabelle der Nachfolgermengen

Mit der Umsetzung dieser Aktion ist der zweite Schritt beendet. Die Abbildung zeigt bereits deutlich, dass Gegenstand X_1 besonders beliebt ist, während Gegenstand X_3 auffallend schlecht beurteilt worden ist. Dieses Ergebnis deckt sich mit den Ausgangsdaten. Ein Informationsverlust ist nicht zu beobachten.

Eine solche Erkenntnis ist bei größeren Datensätzen anhand der Kreuztabelle allerdings meist nur schwer ersichtlich. Ergebnisse der formalen Begriffsanalyse werden daher häufig in Liniendiagrammen dargestellt, da durch diesen Diagrammtyp die volle Information des Kontextes visualisiert und ein struktureller Einblick sichtbar wird. Selbst Implikationen sind ablesbar und Hierarchien werden ersichtlich. Die Schritte drei und vier bestehen folglich aus der Erstellung und Interpretation des zugehörigen Liniendiagramms, wobei aufgrund der trivialen Struktur und der oben bereits beschriebenen Beobachtung auf eine abschließende Interpretation verzichtet werden kann.

Die Erstellung eines Liniendiagramms ist häufig (vor allem bei einem großen Umfang) ein sehr mühevolles Unterfangen. Eine strukturelle Vorgehensweise ist daher dringend erforderlich. Hilfreich sind zudem folgende Hinweise aus der Literatur:

- *„Parallelogrammregel: Bilde Parallelogramme*
- *Geradenregel: Bilde möglichst lange Geraden*
- *Positionierungsregel: gilt c(i)<c(j), dann sollte c(i) auf der Zeichenebene unterhalb von c(j) liegen*
- *Vermeide Kantenkreuzungen*
- *Zeichne in Schichten"* (PETERSEN O.J., S. 3)

Im Fall des Beispiels ist die Erstellung des zugehörigen Liniendiagramms allerdings eine leichte Übung, da der Umfang recht gering ausfällt und das Diagramm linear verläuft.

Abbildung 4.25 Liniendiagramm zum Beispieldatensatz

4.5 Vorgehensweise zur Beseitigung von missing values

Da Probanden einer empirischen Erhebung häufig nicht alle zu bewertenden Gegenstände (in dieser Arbeit sind es Fernsehsendungen) bekannt sind oder Antworten auf persönliche und intrusive Fragen *explizit* verweigert werden, weisen diejenigen Datenmatrizen, die Ausgangspunkt späterer Datenanalysen sind, in der Regel *fehlende Werte* auf. Diese sogenannten *missing values* müssen daher zunächst mit Hilfe eines adäquaten Verfahrens beseitigt werden, weil sich die multivariaten Analyseverfahren, wie die bereits ausführlich beleuchtete Faktorenanalyse, grundsätzlich nur auf vollständiges Datenmaterial anwenden lassen (vgl. RUNTE o.J.). Fehlen nun bei einer empirischen Untersuchung Daten, ist es daher erforderlich, sich mit den Ursachen, der Wirkung und dem Umgang solcher *missing values* zu beschäftigen, um aus einem lückenhaften einen vollständigen Datensatz zu generieren, damit die bekannten Analyseverfahren eingesetzt werden können.

Mittlerweile besitzt zwar jede moderne Statistiksoftware verschiedene Verfahren, um diese *fehlenden Werte* mit Leben zu füllen, doch eignet sich auch hier nicht jede (Standard-) Methode für jeden Anwendungsfall. Gerade für den in dieser Arbeit im Vordergrund stehenden Fall eines Datensatzes, welcher auf Bewertungen und daraus resultierenden Rangfolgen besteht, sind die bekannten Verfahren, wie die Schätzung des *fehlenden Werts* durch das arithmetische Mittel der vorhandenen Daten, schlicht unsinnig. Schließlich besitzen kombinierte Werte (Median, arithmetisches Mittel, ...) bei ordinalskalierten Merkmalen (Noten, Bewertungen, ...) keine Aussagekraft, da die Daten in Kategorien vorliegen und nicht von einer stetigen Bewertungsskala ausgegangen werden kann.

Aus diesem Grund liegt der Fokus dieses Abschnitts auf der Präsentation eines eigenen Vorschlags zur Analyse *ordinalskalierter missing values*, so dass geschätzt werden kann, wie ein Proband einen ihm unbekannten Gegenstand bewerten bzw. einordnen würde, wenn der betrachtete Gegenstand dem Probanden bekannt wäre. Das entsprechende Verfahren basiert auf einem Optimierungsproblem aus der mathematischen (statistischen) Spieltheorie, bei welchem die *missing values* so bestimmt werden, dass die höchstmögliche Fehlanpassung minimal wird. Zur Verdeutlichung soll dieser Vorschlag sowohl mathematisch entwickelt als auch an einem praktischen Anwendungsbeispiel erläutert werden. Dabei wird parallel ein in der Tabellenkalkulation Excel via Makroprogrammierung implementiertes Computerprogramm vorgestellt werden, mit dessen Hilfe die *fehlenden Werte* basierend auf dem eigenen Vorschlag mit wenigen Handgriffen nach und nach beseitigt werden können.

Zuvor soll jedoch basierend auf dem Aufsatz „Missing Values: Konzepte und statische Literatur" von Runte (o.J.) der aktuelle Forschungsstand umrissen werden.

4.5.1 Zur allgemeinen Problematik der missing values

Zur allgemeinen Vorstellung der Problematik sollen in diesem Abschnitt kurz die häufigsten Ursachen für *fehlende Werte* dargelegt werden. Anschließend wird der Unterschied zwischen systematischen und unsystematischen Ausfallmechanismen diskutiert, da eben diese Feststellung sich als Grundlage für den Umgang mit den *missing values* herausstellen wird, so dass abschließend kurz die bekannten Standardverfahren beschrieben werden können und die ablehnende Haltung erörtert wird.

4.5.1.1 Ursachen für missing values

Treten bei einer empirischen Untersuchung *missing values* auf, ist es ratsam, sich zunächst mit den möglichen Ursachen zu beschäftigen. Zwar können die Gründe für *fehlende Werte* äußerst vielfältig sein, doch gelingt es Schnell (1986, S. 24-58) insbesondere die folgenden Aspekte hervorzuheben:

- Explizite Verweigerung einzelner Fragen durch die Befragten,
- „Weiß nicht"- Antworten, basierend auf fehlendem Wissen, fehlender Einstellung zur Thematik oder ausbleibende Beantwortungsmotivation des Befragten,
- Fehlerhaftes oder mangelhaftes Untersuchungsdesign,
- Unangemessene Fragen,
- Unaufmerksamkeit eines Beobachters bzw. ausgelassene/vergessene Fragen,
- Unvollständigkeit von Sekundärdaten,
- Codierungs- und Übertragungsfehler der Daten durch den Anwender.

Wenngleich diese Auflistung ursprünglich der Datenerhebungsmethode „Interview" gewidmet war, lassen sich doch sämtliche Aspekte problemlos auch auf fragebogenbasierte Erhebungsmethoden verallgemeinern.

Offensichtlich ist ein Großteil der auftretenden *fehlenden Werte* von den jeweiligen Forschern selbst verursacht (fehlerhaftes oder mangelhaftes Untersuchungsdesign, Antwortverweigerung, Unvollständigkeit von Sekundärdaten) und kann durch eine exaktere Planung empirischer Untersuchungen und eine fehlerfreie Auswertung (Codierungs- und Übertragungsfehler) schon vorab deutlich reduziert werden. Auch die Antwortmotivation des Befragten lässt sich durch eine ansprechende (optische) Fragebogengestaltung, die den Probanden das Gefühl vermittelt, als Subjekt wahr- und ernst genommen zu werden, steigern (vgl. KIRCHHOFF et al. 2006, S. 25).

In solchen Fällen spricht man von einer *umfragebedingten Missing-Data*. Im Gegensatz dazu wird ein vom Probanden verschuldeter lückenhafter Datensatz als *befragtenbedingte Missing-Data* bezeichnet.

4.5.1.2 Missing-Data Mechanism

Die Behandlung und gewünschte Beseitigung dieser *missing values* basiert vor allem auf den aus den Ursachen für das Fehlen der Daten gezogenen Schlüssen. In der Realität werden die Erhebungseinheiten, welche *fehlende Werte* beinhalten, zwar häufig einfach vollständig ignoriert, doch besteht hierbei das Risiko, dass die aus dieser Ignorierung folgende Datenmatrix die wahren Ergebnisse der empirischen Untersuchung verfälscht. Aus diesem Grund ist es notwendig, sich zunächst zu verdeutlichen, ob die *missing values systematisch* oder *unsystematisch* fehlen. Dabei bedeutet ein *systematisches Fehlen*, dass die Daten nicht „zufällig" fehlen, sondern in direktem *systematischen Zusammenhang* zu anderen Daten stehen. Folgerichtig bedeutet *unsystematisches Fehlen*, dass die Daten tatsächlich rein „zufällig" fehlen. Der Fall „zufällig" *fehlender Werte* wird in der Literatur bevorzugt behandelt (vgl. RUBIN 1976).

Die Unterscheidung zwischen „zufällig" und „nicht zufällig" *fehlenden Werten* geht auf RUBIN (1976) zurück. In der Terminologie der Statistik hat sich dazu folgendes Modell entwickelt:

- **MCAR**: „missing completely at random",
- **MAR**: „missing at random",
- **MNAR**: „missing not at random" (SCHNELL; HILL; ESSER 2007, S. 469).

MCAR bedeutet, dass die Daten völlig zufällig fehlen, etwa weil ein Proband, unaufmerksam wie er sein mag, eine oder mehrere Fragen übersehen hat. Die Wahrscheinlichkeit für das Fehlen dieses Wertes hängt also weder von einer anderen bekannten Variablen, noch von der Ausprägung des Merkmals selbst ab. In solchen Fällen ist es möglich, wenn der Datensatz groß genug und die Anzahl der *fehlenden Werte* sich in einem vertretbaren Rahmen bewegt, die *missing values* zu ignorieren und mit einem (etwas) kleineren Datensatz zu arbeiten.

In der Praxis der empirischen Forschung wird MCAR zwar häufig angenommen, vermutlich, weil der Umgang mit fehlenden Werten sich in diesem Fall als besonders einfach erweist, doch ist die völlige Unabhängigkeit von einer möglichen anderen Variablen nur schwerlich zu beweisen. Zwar wird auch bei MAR davon ausgegangen, dass die Daten zufällig fehlen, doch impliziert MAR nach ANDERSON, BASILEVSKY und HUM (1983, S. 416) nur *„die Unabhängigkeit des Fehlens einer Variablen von jeder anderen Variablen für ein gegebenes Individuum"*. Für diesen, in der Literatur am bevorzugt diskutierten Fall, gibt es einige gängige Standardverfahren zur Schätzung, auf die im weiteren Verlauf nur kurz eingegangen werden soll.

Besonders kompliziert wird es, wenn Daten überhaupt nicht zufällig fehlen (MNAR). Die Verfahren zur Simulation dieser Fälle sind äußerst theoretisch und nehmen in der Literatur kaum Platz ein (vgl. BANKHOFER 1995, S. 85).

Um die drei *grundlegenden* Begriffe konkret zu verdeutlichen, betrachten wir ein bekanntes Beispiel von SCHNELL, HILL und ESSER (2007, S. 468): Bei einer empirischen Untersuchung werden die beiden Variablen *Alter* und *Einkommen* von Probanden erfasst. Nun treten bei der Variablen *Einkommen* an verschiedenen Stellen *missing values* auf. Dann liegt

- MCAR vor, wenn die *fehlenden Werte unabhängig von Einkommen und Alter sind.*
- MAR vor, wenn die *fehlenden Werte unabhängig vom Einkommen aber abhängig vom Alter sind.*
- NMAR vor, wenn die *fehlenden Werte vom Einkommen in bestimmten Altersgruppen abhängen.*

Moderne Statistikprogramme haben mittlerweile Verfahren implementiert, mit denen lückenhafte Datensätze auf MCAR, MAR und MNAR untersucht werden können. In SPSS nennt sich dieses Verfahren „Missing Value Analysis" (MVA) und konzentriert sich auf die Muster fehlender Werte, um mögliche Abhängigkeiten aufzudecken. Dabei sollte sich der geneigte Anwender verdeutlichen, dass eine solche Strukturanalyse nur notwendige, nicht aber hinreichende Bedingungen, zur Annahme *unsystematisch fehlender Werte* liefern kann. Schließlich ist es auch SPSS nicht möglich, die Probanden zu kontaktieren und nachzufragen, warum bestimmte Fragen möglicherweise nicht beantwortet wurden. Die persönliche, aktive Auseinandersetzung mit dem Datensatz und der Untersuchungsgestaltung ist also weiterhin dringend erforderlich, um Aussagen über die Herkunft fehlender Werte zu gewinnen.

4.5.1.3 Bekannte Verfahren zur Beseitigung von missing values

Nachdem Gründe und Auswahlmechanismen lückenhafter Datensätze kurz erläutert wurden, soll, ehe das eigene Verfahren zur rekursiven Beseitigung auftretender *missing values* vorgestellt wird, auf die drei bekanntesten Beseitigungsstrategien eingegangen werden. Diese beziehen sich explizit auf *unsystematisch fehlende Werte*.

Eliminierungsverfahren: Unter der Voraussetzung, dass die *fehlenden Werte* völlig zufällig fehlen (MCAR) können unvollständige Erhebungseinheiten aus der Datenmatrix entfernt werden. Dadurch wird das Ergebnis grundsätzlich nicht verfälscht, doch wird der Datensatz mit jedem *missing value* kleiner, so dass ein großer Datensatz bzw. eine geringe Anzahl *fehlender Werte* eine logische Grundvoraussetzungen für den Fallausschluss sind, da sonst das Risiko besteht, den ursprünglichen Datensatz bis zur Unbrauchbarkeit zu reduzieren. Wenn nicht sicher von MCAR ausgegangen werden kann, sollte dieses Verfahren nicht angewendet werden, will man nicht dem Vorwurf der Verzerrung der Ausgangsdaten ausgesetzt werden (vgl. Runte o.J., S. 9 f). In der Studie dieser Arbeit macht ein vollständiges Eliminierungsverfahren ohnehin keinen Sinn, da bei einem solchen Vorgehen **kein** einziger Proband, bzw. **keine** einzige Fernsehsendung übrig bleiben würde.

Parameterschätzung: Das Verfahren der Parameterschätzung kann in der Praxis grundsätzlich sowohl bei MCAR als auch bei MAR eingesetzt werden. Die *fehlenden Werte* werden basierend auf den vorhandenen Daten der entsprechenden Variablen geschätzt. Nicht selten nutzen Anwender dazu das arithmetische Mittel oder die empirische Varianz. Da im Vergleich zur Eliminierung keine Erhebungseinheiten verloren gehen, hilft das Verfahren den Informationsverlust gering zu halten. Allerdings sollte nicht verschwiegen werden, dass die wahre Verteilung häufig verzerrt wird, da bei der Parameterschätzung die wahren Zusammenhänge sowie die tatsächliche Varianz unterschätzt werden (vgl. Runte o.J., S. 13). Für die hier verwendeten Rangdaten eignet sich dieses Verfahren aus den eingangs angeführten Gründen ohnehin nicht.

Imputationsverfahren: Imputationsverfahren eignen sich sowohl bei MCAR als auch bei MAR. Dabei soll die Datenmatrix ebenfalls durch Schätzung der *fehlenden Werte* gefüllt werden, um erneut den Informationsverlust in Grenzen zu halten. Im Unterschied zur obigen Parameterschätzung werden die Daten aber nicht aus den vorhandenen Daten abgeleitet, sondern unabhängig davon bestimmt. Zusätzlich wird zwischen den *einfachen* und den *multiplen* Imputationsverfahren unterschieden.

Typische Beispiele für die *einfachen Imputationen* sind *Verhältnisschätzungen*, das Verwenden einer *Zufallsauswahl* oder ein geeignetes *Expertenranking*. *Verhältnisschätzungen*, also die Schätzung von Werten unter Annahme einer proportionalen Beziehung, basieren in den meisten Fällen auf Kardinalzahlen und sind daher nur begrenzt einsetzbar. Bei der *Zufallsauswahl* wird einem *fehlenden Wert* ein zufälliger Datenwert zugeordnet, allerdings so, dass die Verteilung des Merkmals weiterhin korrekt abgebildet wird. Nachteil dieses Verfahrens ist die benötigte Bedingung MCAR. *Expertenratings* bestehen erwartungsgemäß daraus, dass Experten subjektiv, basierend auf ihren Erfahrungen, die Lücken im Datensatz füllen (vgl. RUNTE o.J., S. 11 f).

Weit verbreitet sind vor allem aber die *multiplen Imputationen* nach RUBIN (1987). SCHNELL, HILL und ESSER (2007, S. 470) beschreiben das Vorgehen so, dass *fehlende Werte* unter Annahme eines Prozesses, der zum Fehlen der Daten geführt hat, geschätzt werden. Dies geschieht mehrfach (meist fünffach) und führt zu mehreren vollständigen Datensätzen. Einschlägige Computerprogramme analysieren diese Datensätze und fügen sie zu einem gemeinsamen finalen Datensatz zusammen (vgl. SCHWAB 1991, S. 4).

4.5.1.4 Zusammenfassung

In diesem kurzen Abschnitt sollte verdeutlicht worden sein, dass die Thematik der *missing values* sehr weitreichend und der Umgang mit den *fehlenden Werten* nicht eindeutig festgelegt ist. Entscheidend für eine erfolgreiche Behandlung von lückenhaften Datensätzen ist vor allem von der Richtigkeit der Annahme darüber abhängig, ob systematische oder unsystematische Ursachen für die *missing values* verantwortlich sind. Allein die Feststellung dessen führt in der Praxis aber häufig zu Problemen und daraus resultierend zu verzerrten Datensätzen. Auch bei der Auswahl einer geeigneten Schätzstrategie ist sich die Literatur, trotz diverser Untersuchungen, nicht einig. Hier ist der Anwender selbst gefordert, die Vor- und Nachteile im Bezug auf die eigene Stichprobe abzuwägen. Wichtig ist dabei, die Datenstruktur sowie die Skalierung der Daten nicht aus den Augen zu verlieren. Gerade im Hinblick auf ordinalskalierte Rangdaten muss auf ein Vakuum *optimaler Verfahren* hingewiesen werden, wie bei der Aufzählung der Standardverfahren deutlich geworden sein dürfte.

4.5.2 Ein Vorschlag zur Beseitigung fehlender Rangdaten

Basierend auf dem bisher Beschriebenen soll nun ein eigener Vorschlag zur Behandlung *fehlender Werte* vorgestellt werden. Grundlage des Verfahrens ist eine empirische Untersuchung, bei der Probanden auf einer beschränkten diskreten Skala Bewertungen vorgenommen haben, die in Rangdaten transformiert wurden. Dabei wird weiterhin davon ausgegangen, dass sich das Fehlen bestimmter Werte *unsystematisch* erklären lässt.

4.5.2.1 Das mathematische Modell

Seien $\left[0,1\right]$ das reelle Einheitsintervall, $\mathbb{N}$ die Menge der von 0 verschiedenen natürlichen Zahlen und $\left\{\Omega_n\right\}_{n\in\mathbb{N}}$ eine Familie nicht leerer endlicher Mengen, die die Eigenschaft hat, dass zu allen $n\in\mathbb{N}$ ein $k\in\mathbb{N}$ existiert, so dass $\mathbf{\Omega}_n$ eine echte Teilmenge von $\mathbf{\Omega}_k$ ist. Letzteres wird formal durch die Beziehung $\mathbf{\Omega}_n \subset \mathbf{\Omega}_k$ ausgedrückt.

Für jedes $n\in\mathbb{N}$ und jede nicht leere Teilmenge T_n von $\mathbf{\Omega}_n$ heißt eine auf T_n definierte Funktion X *zulässig*, wenn ihr Wertebereich in $\mathbb{N}$ liegt und darüber hinaus eine natürliche Zahl $o(X)$ existiert, so dass $X\left(T_n\right)=\left\{1,\ldots,o\left(X\right)\right\}$ gilt. Die Menge der auf T_n definierten zulässigen Funktionen sei mit $Z\left(T_n\right)$ abgekürzt.

Eine Funktion $\rho_{T_n}: Z\left(T_n\right)\times Z\left(T_n\right)\mapsto\left[0,1\right]$ heißt *Proximität*, wenn sie für alle Funktionen X,Y und Z aus $Z\left(T_n\right)$ folgende Bedingungen erfüllt:

P1: $\rho_{T_n}\left(X,X\right)=1$

P2: $\rho_{T_n}\left(X,Y\right)=\rho_{T_n}\left(Y,X\right)$

P3: $\rho_{T_n}\left(X,Y\right)=1\Rightarrow\rho_{T_n}\left(X,Z\right)=\rho_{T_n}\left(Y,Z\right)$

Wir gehen im Folgenden davon aus, dass für jede mögliche nicht leere Menge T_n genau eine Proximität $\rho_{T_n}: Z\left(T_n\right)\times Z\left(T_n\right)\mapsto\left[0,1\right]$ fest vorgegeben ist.

Seien nun $\mathbf{\Omega}_n\subseteq\mathbf{\Omega}_k$ und $T_n\subseteq\mathbf{\Omega}_n$ sowie $T_k\subseteq\mathbf{\Omega}_n$ nicht leere Mengen, für die $T_n\subset T_k$ gilt. Dann induziert jede Funktion $P\in Z\left(T_k\right)$ in kanonischer Weise eine Funktion $P_{T_n}\in Z\left(T_n\right)$, die rekursiv wie folgt definiert ist:

Zunächst sei $P_{T_n}\left(\omega\right):=1$, falls für alle $\omega'\in T_n$ die Ungleichung $P\left(\omega\right)\le P\left(\omega'\right)$ erfüllt ist. Dann setzen wir $T_n^1:=\left\{\omega\in T_n \mid P_{T_n}\left(\omega\right)\le 1\right\}$.
Seien nun die natürlichen Zahlen $1\le t$ und die Menge $T_n^t:=\left\{\omega\in T_n \mid P_{T_n}\left(\omega\right)\le t\right\}$ gegeben. Dann unterscheiden wir zwei Fälle:

1. Fall: $T_n^t = T_n$.

In diesem Fall ist P_{T_n} bereits vollständig definiert.

2. Fall: $T_n^t \subseteq T_n$.

Für jedes $\omega \in T_n \setminus T_n^t$ definieren wir nun $P_{T_n}(\omega) := t+1$, falls für alle $\omega' \in T_n \setminus T_{\leq n}^t$ die Ungleichung $X(\omega) \leq P(\omega')$ gilt.

Anschließend setzten wir $T_n^{t+1} := \left\{\omega \in T_n \mid P_{T_n}(\omega) \leq t+1\right\}$.

Seien im Folgenden endlich viele nicht notwendig paarweise verschiedene nicht leere Mengen $\Omega_1, \ldots, \Omega_N$ so vorgegeben, dass zu jeder Menge $\Omega_i\ (1 \leq i \leq N)$, die echt in $\Omega := \bigcup_{n=1}^{N} \Omega_n$ enthalten ist, eine Menge $\Omega_k\ (1 \leq k \leq N)$ existiert, die Ω_i echt enthält. Dann betrachten wir für jede natürliche Zahl $n \in \{1, \ldots, N\}$ genau eine fest vorgegebene Funktion $P_n \in Z(\Omega_n)$ und setzen voraus, dass diese Funktionen paarweise verschieden sind.

Für jede natürliche Zahl $1 \leq n \leq N$ interpretieren wir die Menge $\Omega \setminus \Omega_n$ als die Menge der *missing values* von P_n. Unser Ziel ist es nun jede einzelne Funktion $P_n\ (1 \leq n \leq N)$ so zu einer „geeigneten" Funktion $Y_n \in Z(\Omega)$ zu erweitern, dass $Y_{n\Omega_n} = P_n$ gilt. „Geeignet" meint, dass wir uns an den Kriterien „*Permutationsunabhängigkeit*" (Invarianz des Verfahrens gegenüber der Indizierung der Ausgangsfunktionen), „*maximale Ähnlichkeit*" und „*minimale höchste Fehlanpassung*" orientieren.

Betrachten wir zunächst P_1. Dann nehmen wir o.E. an, dass $\Omega_1 \subset \Omega$ gilt. Andernfalls besitzt P_1 keine *missing values* und stimmt mit Y_1 überein. Dann meint „*maximale Ähnlichkeit*", dass wir im ersten Schritt die Menge $M(P_1)$ aller Funktionen $P_k\ (2 \leq k \leq N)$ betrachten, für die $\Omega_1 \subset \Omega_k$ gilt, um anschließend aus $M(P_1)$ die Teilmenge $Max(P_1) := \left\{P_t \in M_1(P_1) \mid \rho_{\Omega_1}(P_1, P_{t\Omega_1}) = \max_{P_k \in M_1(P_1)} \rho_{\Omega_1}(P_1, P_{k\Omega_1})\right\}$ auszuwählen.

Seien nun $\omega_1, \omega_2, \ldots, \omega_s$ diejenigen in $\Omega \setminus \Omega_1$ erhaltenen „Werte", die die Eigenschaft haben, dass zu jedem „Wert" $\omega_g\ (1 \leq g \leq s)$ mindestens eine Funktion $P_t \in Max(P_1)$ existiert, für die $\omega_g \in \Omega_t$ gilt. Dann konzentrieren wir uns zunächst auf ω_1. Sei daher $\Gamma_1 := \Omega_1 \cup \{\omega_1\}$. Da wir sowohl unverbundene als auch verbundene Ränge zulassen, erweitern wir P_1 für jede natürliche Zahl $1 \leq r \leq o(P_1)+1$ zu Funktionen $P_{1r}^u \in Z(\Gamma_1)$ bzw. $P_{1r}^v \in Z(\Gamma_1)$, indem wir für alle $\omega \in \Gamma_1$

$$P_{1r}^{u}\left(\omega\right) := \begin{cases} P_1\left(\omega\right), \text{ falls } \omega \in \Omega_1 \text{ und } P_1\left(\omega\right) \geq r \\ r, \text{ falls } \omega = \omega_1 \\ P_1\left(\omega\right)+1, \text{ falls } \omega \in \Omega_1 \text{ und } P_1\left(\omega\right) \geq r \end{cases}$$

für alle $1 \leq r \leq o\left(P_1\right)+1$

bzw.

$$P_{1r}^{v}\left(\omega\right) := \begin{cases} P_1\left(\omega\right), \text{ falls } \omega \in \Omega_1 \\ r, \text{ falls } \omega = \omega_1 \end{cases}$$

für alle $1 \leq r \leq o\left(P_1\right)$ und $P_{1o\left(P_1\right)+1}^{v} = P_{1o\left(P_1\right)+1}^{u}$ setzen.

Jetzt wählen wir aus $Max\left(P_1\right)$ die Teilmenge $M_1\left(P_1\right)$ aller Funktionen $P_t \in Max\left(P_1\right)$ aus, in deren Definitionsbereich Ω_t die Menge Γ_1 enthalten ist, um anschließend die Ränge $q\left(1 \leq q \leq o\left(P_1\right)+1\right)$ zu bestimmen, für die

$$\min_{P_t \in M_1\left(P_1\right)} \left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\} =$$
$$\max_{1 \leq r \leq o\left(P_1\right)+1} \min_{P_t \in M_1\left(P_1\right)} \left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\}$$

gilt.

Für jeden Rang $1 \leq r \leq o\left(P_1\right)+1$ *wird die „Fehlanpassung" von* P_1 *an eine Funktion* $P_t \in M_1\left(P_1\right)$ *durch* $1-\min\left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\}$ *gemessen. Für jeden Rang* $1 \leq r \leq o\left(P_1\right)+1$ *ist die größtmögliche Fehlanpassung von* P_1 *an die Funktion* $P_t \in M\left(P_1\right)$ *also durch* $\max_{P_t \in M_1\left(P_1\right)} 1-\min\left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\}$ *gegeben. Unser Ziel, den Rang für* ω_1 *in Bezug zu* P_1 *so zu wählen, dass die über alle Ränge* $1 \leq r \leq o\left(P_1\right)+1$ *größtmögliche Fehlanpassung minimal wird, bedeutet daher, dass wir alle Ränge* $p\left(1 \leq p \leq o\left(P_1\right)+1\right)$ *bestimmen müssen, für die*

$$\max_{P_t \in M_1\left(P_1\right)} 1-\min\left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\} =$$
$$\min_{1 \leq r \leq o\left(P_1\right)+1} \max_{P_t \in M_1\left(P_1\right)} 1-\min\left\{\rho_{\Gamma_1}\left(P_{1q}^{u}, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1q}^{v}, P_{t\Gamma_1}\right)\right\}$$

gilt. Dieses Problem ist ein bekanntes, der mathematischen (statistischen) Spieltheorie zugehöriges, Optimierungsproblem. Die Dualität dieses Problems zu obigem erstgenannten Optimierungsproblem beweist die Äquivalenz beider Probleme.

Zur Lösung des erstgenannten Optimierungsproblems bestimmen wir nacheinander die Werte

$$t_1^1 := \min_{P_t \in M_1(P_1)} \left\{ \rho_{\Gamma_1}\left(P_{11}^u, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{11}^v, P_{t\Gamma_1}\right) \right\}$$

$$t_2^1 := \min_{P_t \in M_1(P_1)} \left\{ \rho_{\Gamma_1}\left(P_{12}^u, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{12}^v, P_{t\Gamma_1}\right) \right\}$$

$$\vdots$$

$$t_{o(P_1)+1}^1 := \min_{P_t \in M_1(P_1)} \left\{ \rho_{\Gamma_1}\left(P_{1o(P_1)+1}^u, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1o(P_1)+1}^v, P_{t\Gamma_1}\right) \right\} .$$

Da $M_1\left(P_1\right)$ eine endliche Menge ist, ist die Bestimmung der „Werte” $t_j^1 \left(1 \le j \le o\left(P_1\right)+1\right)$ stets durch vollständiges Durchlaufen der Funktionen aus $M_1\left(P_1\right)$ möglich.

Unter den gewonnenen „Werten” $t_1^1, t_2^1, \ldots, t_{o(P_1)+1}^1$ wählen wir den *besten* Wert t_m^1 aus. Dem Wert t_m^1 entsprechen all diejenigen Ränge $1 \le r \le o\left(P_1\right)+1$, für die $t_m^1 = \min_{P_t \in M_1(P_1)} \left\{ \rho_{\Gamma_1}\left(P_{1r}^u, P_{t\Gamma_1}\right), \rho_{\Gamma_1}\left(P_{1r}^v, P_{t\Gamma_1}\right) \right\}$ gilt. In der Menge dieser Ränge gibt es einen *besten* Rang q_1. Diesen Rang halten wir zunächst fest. Dann setzen wir $\Gamma_2 := \Omega_1 \cup \left\{\omega_2\right\}$ und wählen aus $Max\left(P_1\right)$ die Teilmenge $M_2\left(P_1\right)$ aller Funktionen P_t aus, deren Definitionsbereiche Ω_t die Menge Γ_2 umfasst, um anschließend in der oben beschriebenen Weise minimale „Werte” $t_1^2, t_2^2, \ldots, t_{o(P_1)+1}^2$, einen größten „Wert” t_m^2 und einen zugehörigen größten Rang q_2 zu ermitteln. Mit Hilfe der auf diese Weise Schritt für Schritt gewonnenen Ränge $q_1, q_2, \ldots, q_s$ kann die erste Erweiterung $X_1^1 := \Omega_1^1 := \Omega_1 \cup \left\{\omega_i \mid 1 \le i \le s\right\} \mapsto \mathbb{N}$ dann rekursiv ermittelt werden.

Wegen der bisherigen Permutationsunabhängigkeit des Verfahrens dürfen wir voraussetzen, dass die ermittelten Ränge der *Größe* nach geordnet werden können. Sowohl der *größte* als auch der *kleinste* Rang kann der, je nach gewählter Definition der Ausgangsdaten, *beste* Rang sein. Für den ersten Fall (Fall 2 folgt analog mit umgekehrter Sortierung und wird im nachfolgenden Anwendungsbeispiel erörtert) gilt also $q_1 \le q_2 \le \ldots \le q_s$. Dann existieren natürliche Zahlen $1 \le w_1 < w_2 < \ldots < w_l = s$ mit $q_i = q_1$ für alle $1 \le i \le w_1$, $q_i = q_{w_2}$ für alle $w_1 < i \le w_2$, ..., $q_i = q_s$ für alle $w_l < i \le s$. Nun definieren wir nacheinander $\Omega_1^{w_1} := \Omega_1 \cup \left\{\omega_i \mid 1 \le i \le w_1\right\}$, $\Omega_1^{w_2} := \Omega_1 \cup \left\{\omega_i \mid 1 \le i \le w_2\right\}$, ..., $\Omega_1^{w_l} := \Omega_1 \cup \left\{\omega_i \mid 1 \le i \le s\right\} = \Omega_1^1$ und setzen $M\left(w_1\right) := \left\{1, \ldots, \omega_1\right\}$. Anschließend teilen wir diese Menge in die Mengen

$M^u\left(w_1\right):=$

$\left\{\omega_i \in M\left(w_1\right) \mid \text{ex. keine Funktion } P_{p_i}^v \text{ mit } t_m^i = \rho_{\Gamma_i}\left(P_{p_i}^v, P_{t\Gamma_i}\right) \text{ für eine Funktion } \mathrm{P}_t \in M_i\left(P_1\right)\right\}$

und

$M^v\left(w_1\right):=$

$\left\{\omega_i \in M\left(w_1\right) \mid \text{ex. eine Funktion } P_{p_i}^v \text{ mit } t_m^i = \rho_{\Gamma_i}\left(P_{p_i}^v, P_{t\Gamma_i}\right) \text{ für eine Funktion } P_t \in M_i\left(P_1\right)\right\}$
auf.

Dann definieren wir die Funktion $P_1^{w_1} := \mathbf{\Omega}_1^{w_1} \mapsto \mathbb{N}$ durch

$$P_1^{w_1}\left(\omega\right) := \begin{cases} P_1\left(\omega\right), \text{ falls } \omega \in \mathbf{\Omega}_1 \text{ und } P_1\left(\omega\right) < p_{w_1} \\ p_{w_1}, \text{ falls } \omega \in M^u\left(w_1\right) \\ p_{w_1} \pm 1, \text{ falls } \omega \in M^v\left(w_1\right) \text{ mit "+" in Fall 1 und "-" in Fall 2} \\ P_1\left(\omega\right)+1, \text{ falls } \omega \in \mathbf{\Omega}_1 \text{ und } p_{w_1} \leq P_1\left(\omega\right) \end{cases}$$

` falls $M^u\left(w_1\right) \neq \varnothing$ und $M^v\left(w_1\right) \neq \varnothing$.

$$P_1^{w_1}\left(\omega\right) := \begin{cases} P_1\left(\omega\right), \text{ falls } \omega \in \mathbf{\Omega}_1 \text{ und } P_1\left(\omega\right) < p_{w_1} \\ p_{w_1}, \text{ falls } \omega \in M^u\left(w_1\right) \\ P_1\left(\omega\right)+1, \text{ falls } \omega \in \mathbf{\Omega}_1 \text{ und } p_{w_1} \leq P_1\left(\omega\right) \end{cases}$$

falls $M^u\left(w_1\right) \neq \varnothing$ und $M^v\left(w_1\right) = \varnothing$.

$$P_1^{w_1}\left(\omega\right) := \begin{cases} P_1\left(\omega\right), \text{ falls } \omega \in \mathbf{\Omega}_1 \\ p_{w_1}, \text{ falls } \omega \in M^v\left(w_1\right) \end{cases}$$

falls $M^u\left(w_1\right) = \varnothing$ und $M^v\left(w_1\right) \neq \varnothing$.

Sei nun $1 \leq v < l$. Dann betrachten wir die natürliche Zahl $p_{w_v} + h_v \left(0 \leq h_v, h_v \in \mathbb{N}\right) :=$ $\max\left(\left\{P_1^{w_v}\left(\omega_i\right) \mid 1 \leq i \leq w_v\right\} \cup \left\{P_1^{w_v}\left(\omega\right) \mid \omega \in \mathbf{\Omega}_1 \wedge P_1\left(\omega\right) < p_{v+1}\right\}\right)$ und definieren zunächst $M\left(w_1\right) := \left\{w_v + 1, \ldots, w_{v+1}\right\}$ und dann

$M^u\left(w_{v+1}\right) :=$

$\left\{\omega_i \in M\left(w_{v+1}\right) \mid \text{ex. keine Funktion } P^v_{p_i} \text{ mit } t^i_m = \rho_{\Gamma_i}\left(P^v_{p_i}, P_{t\Gamma_i}\right) \text{ für eine Funktion } P_t \in M_i\left(P_1\right)\right\}$

und

$M^v\left(w_{v+1}\right) :=$

$\left\{\omega_i \in M\left(w_{v+1}\right) \mid \text{ex. eine Funktion } P^v_{p_i} \text{ mit } t^i_m = \rho_{\Gamma_i}\left(P^v_{p_i}, X_{t\Gamma_i}\right) \text{ für eine Funktion } P_t \in M_i\left(P_1\right)\right\}.$

Dann definieren wir die Funktion $P_1^{w_1} := \Omega_1^{w_1} \mapsto \mathbb{N}$ durch

$$P_1^{w_{v+1}}\left(\omega\right) := \begin{cases} P_1^{w_v}\left(\omega\right), \text{ falls } \omega \in \Omega_1^{w_v} \text{ und } P_1^{w_v}\left(\omega\right) < p_{w_{v+1}} + h_v \\ p_{w_{v+1}} + h_v, \text{ falls } \omega \in M^u\left(w_{v+1}\right) \\ p_{w_{v+1}} + h_v \pm 1, \text{ falls } \omega \in M^v\left(w_1\right) \text{ mit "+" in Fall 1 und "-" in Fall 2} \\ P_1^{w_v}\left(\omega\right) + 1, \text{ falls } \omega \in \Omega_1^{w_v} \text{ und } p_{w_{v+1}} + h_v \le P_1^{w_v}\left(\omega\right) \end{cases}$$

falls $M^u\left(w_{v+1}\right) \neq \varnothing$ und $M^v\left(w_{v+1}\right) \neq \varnothing$.

$$P_1^{w_{v+1}}\left(\omega\right) := \begin{cases} P_1^{w_v}\left(\omega\right), \text{ falls } \omega \in \Omega_1^{w_v} \text{ und } P_1^{w_v}\left(\omega\right) < p_{w_{v+1}} + h_v \\ p_{w_{v+1}} + h_v, \text{ falls } \omega \in M^u\left(w_{v+1}\right) \\ P_1^{w_v}\left(\omega\right) + 1, \text{ falls } \omega \in \Omega_1^{w_v} \text{ und } p_{w_{v+1}} + h_v \le P_1^{w_v}\left(\omega\right) \end{cases}$$

falls $M^u\left(w_{v+1}\right) \neq \varnothing$ und $M^v\left(w_{v+1}\right) = \varnothing$

$$P_1^{w_{v+1}}\left(\omega\right) := \begin{cases} P_1^{w_v}\left(\omega\right), \text{ falls } \omega \in \Omega_1^{w_v} \\ p_{w_{v+1}} + h_v, \text{ falls } \omega \in M^v\left(w_{v+1}\right) \end{cases}$$

falls $M^u\left(w_{v+1}\right) = \varnothing$ und $M^v\left(w_{v+1}\right) \neq \varnothing$.

P_1^1 stimmt dann mit $P_1^{w_l}$ überein.

Falls $\Omega_2 \subset \Omega$ fahren wir mit der Menge $M\left(P_2\right) := \left\{X_k \mid 1 \le k \neq 2 \le N \wedge \Omega_2 \subset \Omega_k\right\}$ fort und konstruieren analog zu der beschriebenen Vorgehensweise eine Funktion $P_2^1 : \Omega_2^1 \mapsto \mathbb{N}$. Der erste Schritt ist beendet, wenn alle Funktionen $P_1^1, P_2^1, \ldots, P_N^1$ bestimmt sind. Der zweite

Schritt beginnt dann mit den Mengen $\Omega_1^1, \Omega_2^1, ..., \Omega_N^1$ und den Elementen $P_1^1, P_2^1, ..., P_N^1$ als Ausgangsdaten, um anschließend, genau so wie im ersten Schritt fortgesetzt zu werden.

Auf diese Weise fortfahrend erreichen wir schließlich eine natürliche Zahl d, für die $\Omega_1^d, \Omega_2^d, ..., \Omega_N^d$ gilt. Sobald d erreicht worden ist, sind alle *missing values* beseitigt und die Funktionen $Y_1, Y_2, ..., Y_N$ bestimmt.

Die Beschreibung obiger Vorgehensweise impliziert sofort ihre Permutationsunabhängigkeit.

4.5.2.2 Ein praktisches Anwendungsbeispiel des Verfahrens

Mag das mathematische Modell für die geneigte Leserin und den geneigten Leser auf den ersten Blick noch recht kompliziert und abstrakt wirken, so wird an einem praktischen Beispiel und durch die zur Verdeutlichung eingesetzten Grafiken das logische Vorgehen zur Beseitigung der *fehlenden Rangwerte* sicherlich schnell ersichtlich. Die weiteren Erläuterungen entsprechen explizit der Arbeitsweise des speziell zur Schätzung der *missing values* generierten Computerprogramms in Excel und orientieren sich an den drei großen Arbeitsschritten (1) maximale Ähnlichkeit, (2) minimal höchste Fehlanpassung und (3) rekursives Einsortieren.

Dazu betrachten wir einen Ausgangsdatensatz, der vom Aufbau her dem späteren Untersuchungsdatensatz nachempfunden ist. Wir nehmen an, 25 Probanden $\left(P_1, P_2, ..., P_{25}\right)$ würden gebeten 15 Fernsehsendungen $\left(S_1, S_2, ..., S_{15}\right)$ in ihre jeweilige individuelle Prioritätsordnung zu überführen bzw. der subjektiven Beliebtheit nach zu ordnen.

Dieses Vorgehen entspricht einer Abstimmung auf einer diskreten ordinalskalierten Bewertungsskala, so wie in dem Abschnitt über die mathematischen Grundlagen der Faktorenanalyse eingeführt. Die Rangordnung soll lückenfrei sein. Verbundene Ränge sind zugelassen.

Es sei „1” die *bestmögliche* und „15” die *schlechteste* individuelle Probandenbewertung einer TV-Sendung. Die Werte „25” und „15” sind zufällig gewählt und spielen für die weiteren Ausführungen keine bedeutende Rolle.

Dabei sei folgender Datensatz **A1** entstanden:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	P_{21}	P_{22}	P_{23}	P_{24}	P_{25}
S_1	6	9	1	4	6	4	9	5	13	4	1	9	14	12	4	7	7	2	8	1	1	1	3	4	1
S_2	8	10	13	7		3	10	8	5	8	3	10	4	7	11	2	6	4	2	8	4		4	5	9
S_3	13	15	12	15	11		13	13	15	13	13	14	3	1	15	8	10	3		11	12		5	9	10
S_4	3	11	7	1	9	2	12	3	8	5	2	7			14	6			8	2	2	6	1	13	2
S_5		14	9	3	4		4	14	7	12	5	4	1		12	8	13	5		12	6	7		14	
S_6	4	5	8	13	2		11	6	11	1	6	1	6	3	5	2			5	4	7		1		2
S_7	9	8	11	12		6	7	9	6	6	7	5	7	6	7	4	2	6	6	13	13	5	10	11	8
S_8	2	12	10	8	10	7	8	2	14	7	15	11	11	5	13	8	5		8	14		4		2	1
S_9		2	6	6	5	8	5	5	4	3	8	13	8	4	3	1	12	6	7	3	8	4	9	6	6
S_{10}	12	3	2	2	7	9	2	10	1	9	9	8	9	9	9	5	9	6	3	7	9	3	2	1	5
S_{11}	5	7	14	10				3	10	10	14	2	12	11	8	3	8			5	10	10		12	7
S_{12}	10	13	3	9	3	5	6	13	2	12	10	12	5	8	2	4	4	6	1	9	5		8	3	4
S_{13}	1	1	4	5	1	1	3	1	12	2	11	3	10	2	1	8	3	6	4	6	3	8	11	10	11
S_{14}	11	6	15	11	8		1	9	9	11	12	6	2		6	8	1	6	9	15	14	2	6	7	5
S_{15}	7	4	5	14				8	3		4	11	13	10	10	1	11	1		10	11	9	7	8	3

Abbildung 4.26 Datenmatrix A1: Der lückenhafte Beispieldatensatz

Es ist zu erkennen, dass in der Datenmatrix **A1** einige Werte fehlen (*orange* markiert), weil den entsprechenden Probanden (*zufälligerweise*) einige der 15 Fernsehsendungen unbekannt sind, so dass sie diese nicht bewerteten bzw. einordnen können. Um später mit Methoden der Faktorenanalyse (z.B. Hauptkomponentenmethode mit anschließender Varimax-Rotation) arbeiten zu können, ist ein derart lückenhafter Datensatz, wie eingangs erläutert, vollkommen ungeeignet, so dass diese *missing values* entsprechend der bereits mathematisch beschriebenen Vorgehensweise geschätzt werden sollen. Ziel ist es folglich, zu berechnen, wie ein Proband eine ihm unbekannte Sendung bewertet bzw. eingeordnet hätte.

Im Weiteren kann auf der Basis der wahrscheinlichen Ursachen für die fehlenden Werte (aufgrund von Anonymität können Gründe der absichtlichen Verweigerung wohl ausgeschlossen werden) sowie der in SPSS durchgeführten MVA von einem unsystematischen Fehlen der Daten ausgegangen werden.

Das Verfahren zur Beseitigung der *missing values* kann nun mit Proband P_1 (*grün* markiert) beginnen. P_1 besitzt zwei *fehlende Werte* bei den Sendungen S_5 und S_9, so dass die Grundvoraussetzung $\Omega_1 \subset \Omega$ (nämlich das Vorhandensein *fehlender Werte*) erfüllt ist.

Aufgrund der Permutationsunabhängigkeit des Verfahrens ist es allerdings grundsätzlich irrelevant, mit welchem Probanden begonnen wird. Im Sinne der Übersichtlichkeit empfiehlt sich jedoch eine der Sortierung entsprechende Abhandlung der Probanden.

(1) Maximale Ähnlichkeit:

Um zu schätzen, wie ein Proband (*hier:* P_1) eine ihm nicht bekannte Sendung bewerten bzw. einordnen würde, versuchen wir zunächst einen oder mehrere andere Probanden zu finden, der/die bei den übrigen Fernsehsendungen ein *maximal ähnliches* Bewertungsverhalten gezeigt hat/haben, aber zusätzlich auch die bei P_1 fehlende(n) Sendung(en) bewertet hat/haben.

Basierend auf dieser (maximal ähnlichen) Rangordnung eines oder mehrerer anderen Probanden sollen dann die *fehlenden Werte* von P_1 geschätzt werden. Die jeweilige Ähnlichkeit wird durch den Korrelationskoeffizienten berechnet.

Um diese *maximale Ähnlichkeit* zu einem oder mehreren anderen Probanden mathematisch zu bestimmen, wird zunächst die echte Teilmenge $M\left(P_1\right)$ gebildet, welche diejenigen Probanden beinhaltet, für die $\Omega_1 \subset \Omega_k$ (hier $k \in \left\{2,3,\ldots,25\right\}$) gilt, da der zu P_1 *maximal ähnliche* Proband mehr als die gleichen TV-Sendungen wie P_1 *selbst* bewertet haben muss. Dies geschieht in dem Computerprogramm per Knopfdruck auf den Button des entsprechend programmierten Makros und liefert folgende Datenmatrix **M1**:

	P_1	P_2	P_3	P_4	P_8	P_9	P_{11}	P_{12}	P_{15}	P_{16}	P_{20}	P_{25}
S_1	6	9	1	4	5	13	1	9	4	7	1	1
S_2	8	10	13	7	8	5	3	10	11	2	8	9
S_3	13	15	12	15	13	15	13	14	15	8	11	10
S_4	3	11	7	1	3	8	2	7	14	6	2	2
S_5		14	9	3	14	7	5	4	12	8	12	
S_6	4	5	8	13	6	11	6	1	5	2	4	2
S_7	9	8	11	12	9	6	7	5	7	4	13	8
S_8	2	12	10	8	2	14	15	11	13	8	14	1
S_9		2	6	6	5	4	8	13	3	1	3	6
S_{10}	12	3	2	2	10	1	9	8	9	5	7	5
S_{11}	5	7	14	10	3	10	14	2	8	3	5	7
S_{12}	10	13	3	9	13	2	10	12	2	4	9	4
S_{13}	1	1	4	5	1	12	11	3	1	8	6	11
S_{14}	11	6	15	11	9	9	12	6	6	8	15	5
S_{15}	7	4	5	14	8	3	4	11	10	1	10	3

Abbildung 4.27 Datenmatrix M1: Die Menge M(P_1)

M1 beinhaltet die Probanden $P_2, P_3, P_4, P_8, P_9, P_{11}, P_{12}, P_{15}, P_{16}, P_{20}$ und P_{25}, welche offensichtlich dieses Teilmengenkriterium erfüllen. Diese elf Probanden bilden fortan die Menge $M\left(P_1\right)$.

Kurz erwähnt werden sollte, dass P_{25} zwar selbst einen fehlenden Wert bei Sendung S_5 besitzt, aber dennoch $\Omega_1 \subset \Omega_{25}$ gilt, da P_{25} die Sendung S_9 bewerten konnte.

Da der zur Bestimmung der *maximalen Ähnlichkeit* zwischen P_1 und einem oder mehreren anderen Probanden aus der soeben bestimmten Menge $M\left(P_1\right)$ einzusetzende Korrelationskoeffizient nur auf vollständige Datenmatrizen angewendet werden kann, reduziert das Computerprogramm im nächsten Schritt die obige Datenmatrix **M1** auf diejenigen Sendungen, die

P_1 bewertet hat. Somit wird der Datensatz logischerweise von sämtlichen *missing values* befreit. Die Datenmatrix **M2** präsentiert das Ergebnis dieses Reduktionsvorgangs.

	P_1	P_2	P_3	P_4	P_8	P_9	P_{11}	P_{12}	P_{15}	P_{16}	P_{20}	P_{25}
S_1	6	8	1	3	4	11	1	8	3	7	1	1
S_2	8	9	11	5	6	4	3	9	10	2	7	8
S_3	13	13	10	13	9	13	11	12	13	8	10	9
S_4	3	10	6	1	3	6	2	6	12	6	2	2
S_6	4	4	7	11	5	9	5	1	4	2	3	2
S_7	9	7	9	10	7	5	6	4	6	4	11	7
S_8	2	11	8	6	2	12	13	10	11	8	12	1
S_{10}	12	2	2	2	8	1	7	7	8	5	6	5
S_{11}	5	6	12	8	3	8	12	2	7	3	4	6
S_{12}	10	12	3	7	9	2	8	11	2	4	8	4
S_{13}	1	1	4	4	1	10	9	3	1	8	5	10
S_{14}	11	5	13	9	7	7	10	5	5	8	13	5
S_{15}	7	3	5	12	6	3	4	10	9	1	9	3

Abbildung 4.28 Datenmatrix M2: Die lückenfreie Menge M(P_1)

Da durch das *Löschen* der Sendungen S_5 und S_9 die dort ursprünglich vorhandenen Bewertungen der Probanden aus $M(P_1)$ wegfallen, werden die Bewertungsfunktionen dieser Probanden in **M2** zusätzlich der neuen Situation angepasst. Dies meint insbesondere, dass die Ränge der übrig gebliebenen Sendungen so vergeben werden, als hätte es S_5 und S_9 *nie gegeben.*

So sieht man beispielsweise, dass die Sendung S_1 bei Proband P_2 durch das Kürzen der beiden Zeilen vom neunten (**M1**) auf den achten Rang (**M2**) vorrückt oder sich Sendung S_3 beim gleichen Probanden sogar von Rang 15 (**M1**) auf Rang 13 (**M2**) verbessert hat.

Mit Hilfe der Datenmatrix **M2** können nun die Korrelationen zwischen der Bewertungsfunktion von P_1 und den Bewertungen der Probanden der Menge $M(P_1)$ bestimmt werden.

Um die im faktorenanalytischen Abschnitt bereits erörterte Problematik um den Korrelationskoeffizienten dabei zu umgehen, soll auch bei der Schätzung der missing values statt der einfachen Ränge $(1,2,\ldots n)$ *auf die nunmehr bekannten 0-1-Daten zurückgegriffen werden.*

In diesem Beispiel entspricht der bestmögliche Rang „1" dann dem Tupel $\underbrace{(1,1,1,\ldots,1,1,1)}_{\text{Anzahl der Sendungen}}$

und der schlechteste Rang „15" dem Tupel $\underbrace{(0,0,0,\ldots,0,0,1)}_{\text{Anzahl der Sendungen}}$.

Aus Gründen der Übersichtlichkeit werden die Ränge in dem Computerprogramm allerdings nur dann in 0-1-Daten transformiert, wenn sie mathematisch bzw. rechnerisch von Bedeutung sind, schließlich bestünde die 0-1 Datenmatrix allein in diesem kleinen Anwendungsbeispiel bereits aus bis zu 16×376 *Zellen.*

Um nun aber ohne Kompatibilitätsprobleme die *Ähnlichkeiten* zwischen P_1 und den Probanden aus $M(P_1)$ mit Hilfe des Korrelationskoeffizienten zu bestimmen, wird **M2** an dieser Stelle erstmals in 0-1-Daten übertragen. Basierend auf dieser 0-1-Datenmatrix errechnet das Computerprogramm dann automatisch die gesuchten Korrelationen und liefert die der Größe des jeweiligen Korrelationskoeffizienten nach geordnete Tabelle **CorrP1**:

Korr(P_1,P_8)	0,745446
Korr(P_1,P_{20})	0,523810
Korr(P_1,P_{12})	0,486997
Korr(P_1,P_2)	0,476190
Korr(P_1,P_4)	0,452381
Korr(P_1,P_{25})	0,440088
Korr(P_1,P_{15})	0,428571
Korr(P_1,P_3)	0,357143
Korr(P_1,P_{11})	0,357143
Korr(P_1,P_{16})	0,295189
Korr(P_1,P_9)	0,166667

Abbildung 4.29 CorrP1: Die maximale Ähnlichkeit

Offensichtlich hat P_8 ein *maximal ähnliches* Abstimmungsverhalten wie P_1 gezeigt. Die Korrelationen zwischen P_1 und den übrigen Probanden weisen dagegen auf wesentlich geringere Ähnlichkeiten hin. P_8 bildet also als einziges Element den Inhalt der Menge $Max(P_1)$ des mathematischen Modells. Diese Menge enthält nur dann mehr als einen Probanden, wenn die größte Korrelation mehrfach auftritt.

(2) Minimale höchste Fehlanpassung:

Nach diesem ersten Schritt ist bekannt, dass die *fehlenden Werte* von Proband P_1 mit Hilfe der Bewertungen von Proband P_8 geschätzt werden sollen. Wie im mathematischen Modell angedeutet, orientiert sich das weitere Vorgehen an einem aus der mathematischen (statistischen) Spieltheorie bekannten Optimierungsproblem: Die *missing values* sollen so bestimmt werden, dass die *größtmögliche Fehlanpassung* letztlich *minimal* wird.

Dazu werden zunächst alle *möglichen Ränge* für die *fehlenden Sendungen* innerhalb der Bewertungsfunktion von P_1 simuliert. Das bedeutet in diesem Fall, dass für die Sendungen S_5

und S_9 alle möglichen Bewertungen jeweils in unverbundener als auch in verbundener Weise getrennt voneinander angegeben und festgehalten werden, um im Anschluss sämtliche Bewertungsalternativen von P_1 mit der Bewertungsfunktion von P_8 zu vergleichen, um den *bestmöglichen* Rang für einen *fehlenden Wert* zu bestimmen.

In betrachtetem Beispiel sind dazu ω_1 und ω_2 diejenigen Werte, die bei P_1 fehlen, nicht aber bei P_8 (also die Bewertungen der Fernsehsendungen S_5 und S_9). In dem Computerprogramm werden diese mit Bezug zu der Nummerierung der gesuchten Sendungen als ω_5 und ω_9 bezeichnet.

Wir betrachten zunächst ω_5 und bilden die Menge $\Gamma_5 := \Omega_1 \cup \omega_5$. Γ_5 ist also diejenige Menge, die die vorhandenen Bewertungen von P_1 mit allen möglichen Bewertungen der bisher fehlenden Sendung S_5 vereint. Anschließend betrachten wir ω_9 und bilden analog die Menge $\Gamma_9 := \Omega_1 \cup \omega_9$. Beide Mengen sind logischerweise identisch und werden durch das Computerprogramm in Form der Datenmatrix **R** illustriert.

	$P^v_{1r(1)}$	$P^u_{1r(1)}$	$P^v_{1r(2)}$	$P^u_{1r(2)}$	$P^v_{1r(3)}$	$P^u_{1r(3)}$	$P^v_{1r(4)}$	$P^u_{1r(4)}$	$P^v_{1r(5)}$	$P^u_{1r(5)}$	$P^v_{1r(6)}$	$P^u_{1r(6)}$	$P^v_{1r(7)}$
S_1	6	7	6	7	6	7	6	7	6	7	6	7	6
S_2	8	9	8	9	8	9	8	9	8	9	8	9	8
S_3	13	14	13	14	13	14	13	14	13	14	13	14	13
S_4	3	4	3	4	3	4	3	3	3	3	3	3	3
S_6	4	5	4	5	4	5	4	5	4	4	4	4	4
S_7	9	10	9	10	9	10	9	10	9	10	9	10	9
S_8	2	3	2	3	2	2	2	2	2	2	2	2	2
S_{10}	12	13	12	13	12	13	12	13	12	13	12	13	12
S_{11}	5	6	5	6	5	6	5	6	5	6	5	5	5
S_{12}	10	11	10	11	10	11	10	11	10	11	10	11	10
S_{13}	1	2	1	1	1	1	1	1	1	1	1	1	1
S_{14}	11	12	11	12	11	12	11	12	11	12	11	12	11
S_{15}	7	8	7	8	7	8	7	8	7	8	7	8	7
S_{neu}	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**	**5**	**5**	**6**	**6**	**7**

Abbildung 4.30 Datenmatrix R: Alle Möglichkeiten eine lückenhafte Stelle zu besetzen

Die linke Spalte zeigt weiterhin die Fernsehsendungen ohne die von P_1 nicht bewerteten Sendungen S_5 und S_9. Hinzugekommen ist dagegen die unterste Zeile „S_{neu}", in der die möglichen Bewertungen für eine der fehlenden Sendungen eingetragen werden.

In der Datenmatrix **R** simuliert das Computerprogramm also alle Möglichkeiten, wie S_{neu} (wahlweise ω_5 oder ω_9) bewertet werden könnte und welchen Einfluss diese zusätzliche Bewertung auf das Ranking der 13 übrigen Sendungen hat.

Für jeden Rang werden sowohl *unverbundene* als auch *verbundene* Ränge zugelassen, so dass auch jeweils beide Möglichkeiten betrachtet werden. Wie im mathematischen Modell beschrieben, verändern sich die übrigen Bewertungen bei einem neuen *verbundenen* Rang nicht, während bei einem neuen *unverbundenen* Rang die Werte, die *schlechter* oder gleich des *neuen* Ranges sind, um 1 addiert werden (mit Ausnahme des schlechtesten Rangs; hier 15). Somit zeigt die Funktion $P^{v}_{1r(3)}$ zum Beispiel, wie die Bewertungen von P_1 aussehen würden, wenn P_1 eine der nicht bewerteten Sendungen mit dem verbundenen Rang 3 bewertet hätte.

Da diese Tabelle grundsätzlich für alle $\omega_1, \omega_2, ..., \omega_s$ identisch ist, wurde bei der Programmierung des Excel-Makros die allgemeingültige Bezeichnung S_{neu} gewählt.

Nachdem bisher lediglich alle möglichen Bewertungsfunktionen von P_1 unter Hinzunahme eines fehlenden Werts festgehalten wurden, soll nun die *bestmögliche* Alternative individuell für jeden *missing value* ausgewählt werden.

Dazu werden zunächst die Probanden aus $Max\left(P_1\right)$ ausgewählt, die ω_5 bzw. ω_9 bewertet haben. Mit Hilfe der Bewertungen dieser Probanden sollen dann jeweils die *bestmöglichen Werte* bei P_1 geschätzt werden. In diesem Beispiel handelt es sich in beiden Fällen natürlich einzig und alleine um Proband P_8, so dass P_8 damit gleichzeitig zum einzigen Element der Mengen $M_5\left(P_1\right)$ und $M_9\left(P_1\right)$ wird. Da P_8 also sowohl ω_5 als auch ω_9 bewertet hat, können sämtliche bei P_1 *fehlenden Werte* in einem Durchlauf berechnet werden.

Zur Erinnerung: Wäre (nur) Proband P_{25} Element der Menge $Max\left(P_1\right)$, wäre es nicht möglich beide missing values in einem Durchlauf zu bestimmen, da P_{25} selbst einen fehlenden Wert aufweist.

Zur Lösung des ersten Optimierungsproblems, dem Bestimmen der *bestmöglichen Ränge*, müssen im nächsten Schritt nun die Werte $t^5_1, t^5_2, ..., t^5_{o\left(P_1\right)+1}$ für Sendung S_5 sowie die Werte $t^9_1, t^9_2, ..., t^9_{o\left(P_1\right)+1}$ für Sendung S_9 bestimmt werden, was mit Hilfe der Tabelle **RW** geschieht.

Zunächst betrachten wir das Vorgehen zur Bestimmung der Werte $t^5_1, t^5_2, ..., t^5_{o\left(P_1\right)+1}$.

Die Tabelle **RW** besteht aus zwei Teilen:

	$P^v_{1r(1)}$	$P^u_{1r(1)}$	$P^v_{1r(2)}$	$P^u_{1r(2)}$	$P^v_{1r(3)}$	$P^u_{1r(3)}$	$P^v_{1r(4)}$	$P^u_{1r(4)}$	$P^v_{1r(5)}$	...	$P^u_{1r(13)}$	$P^v_{1r(14)}$	ω_5
S_1	6	7	6	7	6	7	6	7	6	...	6	6	4
S_2	8	9	8	9	8	9	8	9	8	...	8	8	6
S_3	13	14	13	14	13	14	13	14	13	...	14	13	9
S_4	3	4	3	4	3	4	3	3	3	...	3	3	3
S_6	4	5	4	5	4	5	4	5	4	...	4	4	5
S_7	9	10	9	10	9	10	9	10	9	...	9	9	7
S_8	2	3	2	3	2	2	2	2	2	...	2	2	2
S_{10}	12	13	12	13	12	13	12	13	12	...	12	12	8
S_{11}	5	6	5	6	5	6	5	6	5	...	5	5	3
S_{12}	10	11	10	11	10	11	10	11	10	...	10	10	9
S_{13}	1	2	1	1	1	1	1	1	1	...	1	1	1
S_{14}	11	12	11	12	11	12	11	12	11	...	11	11	7
S_{15}	7	8	7	8	7	8	7	8	7	...	7	7	6
S_{neu}	1	1	2	2	3	3	4	4	5	...	13	14	**10**

Abbildung 4.31 Datenmatrix RW: Alle Möglichkeiten eine lückenhafte Stelle zu besetzen und der Proband maximaler Ähnlichkeit

Sie beinhaltet einerseits die aus der Datenmatrix **R** bekannten Bewertungen von P_1 für die 13 benoteten Sendungen inklusive aller Möglichkeiten, wie P_1 die Sendung S_5 hätte bewerten können.

Anderseits zeigt die Tabelle in der letzten Spalte ω_5 die Bewertungsfunktion von P_8 unter Hinzunahme der entsprechenden Bewertung für Sendung S_5 in der Zeile S_{neu}. Auf die anderen bei P_1 fehlenden Sendungen (hier: S_9) wird dagegen weiterhin verzichtet.

Der Wert t_i^5 $(1 \le i \le o(P_1)+1)$ gibt mathematisch die minimale Korrelation zwischen der Bewertungsfunktion von P_1 (für die 13 bewerteten Sendungen inklusive der Bewertung i für Sendung S_5) und der Bewertungsfunktion von P_8 (ω_5) für den verbundenen Rang i oder für den unverbundenen Rang i als mögliche Bewertung der Fernsehsendung S_5 an. Aus all diesen *minimalen* Werten $t_1^5, t_2^5, \ldots, t_{o(P_1)+1}^5$ wird nun abschließend derjenige Rang ausgewählt und festgehalten, welcher die *maximale* Korrelation aufweist. Dieses i lässt sich inhaltlich nun als der *bestmögliche* Rang für die TV-Sendung S_5 interpretieren (bei mehreren Rängen mit gleicher Korrelation wird der *beste* Rang ausgewählt). Der Algorithmus führt dazu, dass der höchstmögliche Verlust, der entstehen kann, wenn man die durch den *missing value* entstandene Lücke in der Datenmatrix durch das Einfügen eines *falschen* Wertes schließt, minimiert wird. Die Bestimmung der Werte $t_1^9, t_2^9, \ldots, t_{o(P_1)+1}^9$ verläuft analog.

Selbstverständlich müssen die Rangwerte vor der Berechnung der Korrelationen in 0-1-Daten transformiert werden.

In dem betrachteten Beispiel ergibt sich für den fehlenden Wert bei Sendung S_5 der *best-mögliche* Rang 10 (unverbunden) und für den fehlenden Wert bei Sendung S_9 der *best-mögliche* Rang 4 (unverbunden). Beide Ergebnisse werden in der nachfolgenden Grafik festgehalten:

ω_5		ω_9	
$P^u_{1r(10)}$	0,744	$P^u_{1r(4)}$	0,713

ω_5	ω_9
$P^u_{1r(10)}$	$P^u_{1r(4)}$

daraus folgt

q_2	q_1
$P^u_{1r(10)}$	$P^u_{1r(4)}$

und

M^u	M^v
q2	
q1	

q_2	q_1
w_5	w_9
r(10)u	r(4)u

Abbildung 4.32 Die bestmöglichen Ränge

Die gewonnenen Ränge werden nun der Größe nach geordnet und mit q_1 sowie q_2 bezeichnet. Da im Unterschied zum Verfahren im theoretischen Teil, wie angekündigt, hier „1" der *best-möglichen* Bewertung einer Fernsehsendung entspricht, wird q_1 dem *besseren* der beiden Werte (Rang 4) und q_2 dem *schlechteren* der beiden Werte (Rang 10) zugeordnet.

Die beiden Ränge können nun auf die beiden Mengen $M^u\left(w_1\right)$ und $M^v\left(w_1\right)$ verteilt werden. $M^u\left(w_1\right)$ enthält logischerweise alle unverbundenen Ränge, während $M^v\left(w_1\right)$ alle verbundenen Ränge beinhaltet. Wie oben zu sehen ist, führt das Computerprogramm auch diese, im vorliegenden Fall zugegebenermaßen simple Aufgabe, korrekt durch, indem sowohl q_1 als auch q_2 bei $M^u\left(w_1\right)$ einsortiert werden.

(3) Rekursives Einsortieren:

Es ist bis dato gelungen, bestmögliche Ränge für die von Proband P_1 nicht bewerteten Sendungen S_5 und S_9 zu schätzen, bei denen die größtmögliche Fehlanpassung minimal wird. Diese Schätzungen beruhen auf dem zu P_1 maximal ähnlichen Abstimmungsverhalten von Proband P_8. Abschließend gilt es nun, diese gefundenen Ränge in das bisherige Bewertungsmuster von P_1 rekursiv einzusortieren, damit die Lücken im Datensatz endlich geschlossen werden können.

Die Tabelle **S** dokumentiert diesen letzten Schritt:

	P_1
S_1	6
S_2	8
S_3	13
S_4	3
S_5	
S_6	4
S_7	9
S_8	2
S_9	
S_{10}	12
S_{11}	5
S_{12}	10
S_{13}	1
S_{14}	11
S_{15}	7

Schritt 1	Schritt 2
6	7
8	9
14	15
3	3
10	11
4	5
9	10
2	2
	4
13	14
5	6
11	12
1	1
12	13
7	8

P_1^1
7
9
15
3
11
5
10
2
4
14
6
12
1
13
8

Abbildung 4.33 Datenmatrix S: Der rekursive Einsortiervorgang

Wie im mathematischen Modell gezeigt, muss zunächst das geeignete Verfahren ausgewählt werden. Hierzu betrachten wir den Inhalt der Mengen $M^u(w_1)$ und $M^v(w_1)$. In unserem Anwendungsbeispiel stellt sich die Situation so dar, dass $\{q_1, q_2\} \in M^u(w_1)$ und $M^v(w_1) = \varnothing$ gilt. Dementsprechend wendet das Computerprogramm Verfahren 2 aus dem mathematischen Modell an.

Dabei wird zunächst der *schlechteste* Wert (hier: q_2) einsortiert. Hierbei handelt es sich um Rang 10, der bei Sendung S_5 im ersten Schritt („Schritt 1" in **S**) eingefügt wird. Da es sich um einen unverbundenen Rang handelt, werden alle bereits vorhanden Bewertungen ≥ 10 um einen Rang erhöht (*verschlechtert*), während alle bereits vorhandenen Bewertungen < 10 unberührt bleiben. In Schritt 2 („Schritt 2" in **S**) wird auf die gleiche Art und Weise der unverbundene Rang 4 bei Sendung S_9 einsortiert und wir erhalten den *missing value*-freien Probanden $P_1^1 = Y_1$, so dass wir uns dem nächsten Probanden widmen können.

Wenn an dieser Stelle noch weitere *missing values* auftreten würden, müsste dieses Verfahren für P_1 wiederholt werden, nachdem dieser erste Durchgang für alle anderen Probanden ebenfalls durchgeführt wurde und damit ein neuer Datensatz mit weniger Lücken als Grundlage bestünde.

5. Datengrundlage: die Stichprobe

5.1 Die Altersgruppe und der Bildungshintergrund der Probanden

Nachdem die mathematischen und statistischen Vorüberlegungen abgeschlossen sind, stellt sich zu Beginn eines Forschungsprojekts zunächst die Frage, über wen man Aussagen treffen möchte und folglich wer an der Befragung teilnehmen soll (vgl. Kirchhoff et al. 2006, S. 15). Zunächst ist die Entscheidung getroffen worden, dass sich die Studie auf Schülerinnen und Schüler des 7. Jahrgangs beziehen soll.

Für diese Auswahl spricht, dass die zwölf- bis vierzehnjährigen sich in einem Lebensabschnitt befinden, der von Selbstständigkeit und Unselbstständigkeit sowie von Eigen- und Fremdbestimmung geprägt ist (vgl. Limbourg 2009). Nach Eriksons entwicklungspsychologischem Stufenmodell von 1959 befindet sich der Mensch bekanntlich ab dem 12. Lebensjahr in der Jugend bzw. der Adoleszenz. Havighurst (1948, S. 2 ff) definiert als die wichtigste Entwicklungsaufgabe im Jugendalter den Aufbau einer eigenen, möglichst positiven Identität in Folge der Beantwortung der Fragen *wer man sei*, *was man könne* und *was man wolle.* Sozialisationstheoretisch bedeutet dies, dass die Jugendlichen Handlungskompetenzen erwerben sollen, um produktiv mit den Anforderungen der Umwelt umgehen zu können. Um diese Fähigkeiten zu erlangen, müssen sie lernen, eigene Motive, Interessen und Bedürfnisse zu definieren und einzubringen. Folglich sollen in dieser Lebensphase die Voraussetzungen für ein selbstständiges Handeln in allen gesellschaftlichen Bereichen erworben werden (vgl. Limbourg 2009). Daher geht der Verfasser davon aus, dass die Jugendlichen auch zunehmend selbstbestimmt über ihr Fernseh- und Freizeitverhalten entscheiden. Die in Kapitel 2 vorgestellte JIM-Studie untermauert diese Vermutung und lässt die Siebtklässler als optimale Probandengruppe erscheinen.

Eine Vollerhebung aller Siebtklässler der Bundesrepublik Deutschland oder auch Nordrhein-Westfalens kann natürlich nicht der Anspruch dieser Arbeit sein, wobei es dennoch das ehrgeizige Ziel des Verfassers ist, eine möglichst repräsentative Probandenauswahl zu generieren, schließlich ist bei jeder Studie stets *„darauf zu achten, dass die Stichprobe ein möglichst ge-*

treues Abbild der Gesamtpopulation ist" (FAHRMEIR 1997, S. 14). Da derartige empirische Untersuchungen auch für die teilnehmenden Schulen in einem straffen Stundenplan unplanmäßige Mehrarbeit und den Verzicht auf einige Unterrichtsstunden bedeuten, ist die Auswahl der Probanden natürlich nicht beliebig. Nichtsdestotrotz konnten vier verschiedene Schulen dreier Schulformen für das Projekt begeistert werden, die in ihrer Zusammensetzung durchaus spannende und aussagekräftige Ergebnisse versprechen. Bei den vier Schulen handelt es sich um das Pestalozzi-Gymnasium Unna, die Gesamtschule Kamen, die Gesamtschule Süd in Essen und die Gemeinschaftshauptschule an der Bruchstraße in Mülheim an der Ruhr. Statistisch ausgedrückt stellen folglich die teilnehmenden Schülerinnen und Schüler der jeweiligen siebten Jahrgänge der aufgeführten Schulen die Grundgesamtheit der Umfrage dar. Insgesamt konnten 232 Probanden aus neun verschiedenen Klassen befragt werden.

Es folgt eine kurze Vorstellung der Schulen, die einen Eindruck vermitteln soll, warum diese Schulen sowohl einzeln als auch in der Kombination für die durchgeführte Untersuchung geeignet erscheinen. Aufgrund des Forschungsthemas sollen vor allem die soziale Zusammensetzung, der mathematische Fachbereich und der Umgang mit Neuen Medien im Zentrum der Überlegungen stehen.

5.2 Pestalozzi-Gymnasium Unna

Glaubt man der Meinung vieler aktueller und ehemaliger Schülerinnen und Schüler, ist das nach dem Schweizer Pädagogen JOHANN HEINRICH PESTALOZZI (1746-1827) benannte Pestalozzi-Gymnasium die anspruchsvollste Schule in Unna, einer etwa 66.000 Einwohner zählenden Kreisstadt im östlichen Ruhrgebiet. Die Wurzeln des Gymnasiums reichen bis in das Jahr 1295 zurück, zumindest verweist eine Urkunde des Klosters Himmelpforten auf einen entsprechenden Vorläufer. Offiziell wurde die Schule schließlich am 01. Mai 1905 als erstes Unnaer Gymnasium gegründet, ehe sie 1971 in „Pestalozzi-Gymnasium" (PGU) umbenannt wurde. 1986 beschloss die Stadt Unna die Dreizügigkeit sowie den Umzug des Gymnasiums in das neue Gebäude an der Morgenstraße.

Seitdem konnte die Schule zahlreiche Auszeichnungen sammeln. So kann sich das PGU heutzutage beispielsweise als „Selbstständige Schule" (seit 2002) und „UNESCO-Schule" (seit 2005) bezeichnen. Zudem erhielt die Schule das Qualitätssiegel „Schule => Beruf" (2007) für sein Berufsorientierungskonzept sowie die Gütesiegel „Agenda21" (2008) und „Individuelle Förderung" (2008) (vgl. Schulprogramm des Pestalozzi-Gymnasium Unna).

Aktuell konkurriert das Pestalozzi-Gymnasium mit zwei weiteren Gymnasien um die Gunst der Kinder und Eltern in Unna, verzeichnete dabei aber in den letzten Jahren einen wahren Anmeldungsansturm, der vor allem darauf zurückzuführen ist, dass durch die Attraktivität der Schullandschaft der Anteil der Schülerinnen und Schüler aus anderen Kommunen beständig wächst (vgl. STEPHAN 2010).

In Anlehnung an JOHANN HEINRICH PESTALOZZIS Philosophie werden aktuell 996 Schülerinnen und Schülern von 61 Lehrerinnen und Lehrern sowie acht Referendarinnen und Referendaren nach dem Motto *Lernen mit Kopf, Herz und Hand* und dem Ziel, jedes Kind in den Blick zu nehmen, unterrichtet. Nach wie vor haftet dem PGU dabei der Ruf an, eine elitäre Schülerschaft zu besitzen, so dass es nicht verwundert, dass nur 3,12 % der Schülerinnen und Schüler einen Migrationshintergrund aufweisen (vgl. Kreisstadt Unna, der Bürgermeister 2011). Wenngleich dieser Wert äußerst gering ist, bildet er im Rahmen der empirischen Studie des Verfassers den landesweiten Ausländeranteil von 4,4 % an Gymnasien in Nordrhein-Westfalen hinreichend genau ab (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen 2012). Die Abiturnoten in Zeiten des Zentralabiturs bestätigen ebenfalls die hohe Unterrichtsanforderung und -qualität des Gymnasiums an der Morgenstraße. 2011 erreichten 2 % der Schülerinnen und Schüler die Bestnote 1,0, wobei mit 822 von 840 Punkten sogar ein neuer Bestwert aufgestellt werden konnte. Zudem erhielten insgesamt 30 % der Abiturienten eine 1 vor dem Komma, während weitere 47 % immerhin noch eine 2 vor dem Komma vorweisen konnte (vgl. NEITZEL 2011). Damit liegt das PGU auch in dieser Statistik oberhalb des landesweiten Durchschnitts an Gymnasien (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen 2012).

Hinsichtlich des Umgangs mit Neuen Medien qualifiziert sich das Pestalozzi-Gymnasium ebenfalls für die Teilnahme an dem Forschungsprojekt, da sich die Schule dazu entschieden hat, bereits ab dem fünften Jahrgang wöchentlich eine Doppelstunde „Medienerziehung" durchzuführen. Dadurch werden die Schülerinnen und Schüler im Verlauf der Sekundarstufe I einerseits dazu befähigt, spätestens mit dem Eintritt in die gymnasiale Oberstufe über einen qualifizierten und wissenschaftsadäquaten Umgang mit dem Computer zu verfügen, sowie andererseits zu einer verantwortungsvollen und kritischen Auseinandersetzung mit den neuen visuellen Medien erzogen. Diese Kenntnisse werden zusätzlich durch den fachspezifischen Einsatz der entsprechenden Medien vertieft und weiter gefördert. Ähnlich wie bei den nachfolgenden Schulen ist es aber nicht das Ziel des Gymnasiums, die traditionellen Medien aus dem Klassenzimmer zu verbannen. Daher wird den Schülerinnen und Schüler auch der reflektierte Umgang mit den klassischen Medien, beispielsweise in Form der Schulbibliothek, näher gebracht, so dass der Unterricht zur Symbiose aus klassischen und modernen Medien wird (vgl. Schulprogramm des Pestalozzi-Gymnasium Unna).

Neben dem traditionellen Mathematikunterricht, welcher sich natürlich an den Kernlehrplänen des Landes Nordrhein-Westfalen orientiert, bietet das Pestalozzi-Gymnasium mittlerweile seit über zehn Jahren eine mathematisch geprägte Arbeitsgemeinschaft an, die sich beispielsweise mit Kreuzsummen, magischen Quadraten, Zahlenrätseln, Kryptogrammen, Pentominos und Logik auseinandersetzt und deren Ziel es ist, Spaß an Mathematik und Knobelaufgaben zu vermitteln. Die „Knobel AG" wirbt damit, dass Mathematik weniger Rechnen als vor allem Denken sei. Dass die Teilnahme an der Arbeitsgemeinschaft von Erfolg gekrönt ist, zeigen die Ergebnisse der entsprechenden Schülerinnen und Schüler bei den bundesweiten Mathematikwettbewerben. Nicht selten sind gerade diese Kinder und Jugendliche bei den landesweiten Finalrunden erfolgreich vertreten (vgl. ebd.).

5.3 Gesamtschule Kamen

Die Gesamtschule liegt in der knapp 43.500 Einwohner zählenden Stadt Kamen an der Peripherie des Stadtkerns im Übergang zur ländlichen westfälischen Landschaft. Auch Kamen, zwischen Hamm und Dortmund liegend, war als Teil des östlichen Ruhrgebiets stark von der Bergbauindustrie geprägt. Lange Zeit war die Zeche Monopol das Aushängeschild der Stadt. Mit dem Anwachsen des Bergbaus kam es in Kamen zu einer großen Immigration ausländischer Arbeiter.

1969, zur Zeit der Bildungsreform, galt es die Frage zu beantworten, wie das Schulwesen zu optimieren sei. Letztlich wurde die gefeierte Entscheidung getroffen, eine integrierte Gesamtschule mit gymnasialer Oberstufe zu errichten, die bereits 1975 zu den größten weiterführenden Schulen in ganz Deutschland gehörte. Zwischenzeitlich wurde die Schule aufgrund ihrer ständig wachsenden Schülerzahl sogar in zwei benachbarte Gesamtschulen, die Carlo-Schmid und die Hermann-Ehlers-Gesamtschule, unterteilt. Diese beiden Bildungseinrichtungen führten schnell dazu, dass sich die Zahl der Abiturienten in Kamen verdoppelte. Während die Gesamtschulen ihrem Schulprofil nach und nach immer schärfere Konturen gaben, schulkulturelle Aufführungen, ökologische Initiativen und Projekte gegen Krieg und Rassismus waren die Folge, errichtete die Stadt Kamen Ende der 90er Jahre zusätzlich zu den bestehenden Gesamtschulen und dem Gymnasium noch jeweils eine Real- und eine Hauptschule. Die geteilten Gesamtschulen wurden zur „Gesamtschule Kamen" wiedervereinigt und zu einer Angebotsschule umfunktioniert.

Derzeit besuchen 1.375 Schülerinnen und Schüler die Gesamtschule Kamen, während sich knapp 100 Lehrerinnen und Lehrer sowie zwei Sozialpädagogen für die Ausbildung verantwortlich zeigen. Das Einzugsgebiet der Gesamtschule umfasst sämtliche Stadtteile Kamens (Mitte, Methler, Südkamen, Heeren-Werve, Derne und Rottum) und bildet somit die sozialen Schichten der Stadt hinreichend genau ab. Immerhin liegt der Anteil der Jugendlichen aus Migrationsfamilien mit knapp 15 % ungefähr auf Realschulniveau und nur knapp über dem Ausländeranteil Nordrhein-Westfalens (11 %) (vgl. Statistisches Bundesamt Deutschland o.J.). Gleichzeitig ist die Quote der Eltern, die die Hochschulreife besitzen, mit 31 % relativ hoch.

Dabei handle es sich nach DÖRGER um ein mittleres Sozialschichtniveau, das wiederum mit der Schülerschaft an Realschulen vergleichbar sei (vgl. DÖRGER o.J.). Darüber hinaus durchlaufen etwa 30 % der Abiturienten der Stadt Kamen die Gesamtschule. Wenn man bedenkt, dass „nur” knapp die Hälfte aller Schülerinnen und Schüler Kamens diese Schule besuchen, ist der Wert umso positiver aufzufassen. Zuletzt lag dabei die durchschnittliche Abiturnote (2,75) sogar oberhalb des gesamtschulischen Durchschnitts (2,87) in Nordrhein-Westfalen.

Ein Großteil des Lernens findet an der Gesamtschule in Projekten statt. Auch im Bereich der Neuen Medien hat die Schule erkannt, dass es nicht mehr möglich ist, diese aus dem Leben der Kinder zu verbannen. Aus diesem Grund geht die Schule offensiv mit der Thematik um. Beispielsweise werden schon in der Erprobungsstufe im Gesellschaftslehreunterricht die Bedeutung des Fernsehens und seine Faszination für Kinder und Jugendliche behandelt. Außerdem wird im Unterricht diskutiert, wie Medien sinnvoll zum Lernen genutzt werden können. Darüber hinaus erweiterte die Gesamtschule ihr Medienkonzept zuletzt um einen weiteren Baustein. Schon ab der 5. Klasse sollen die Kinder lernen, mit dem Computerprogramm „Microsoft Word” umzugehen. Im Zusammenhang mit der Einführung der Rechtschreibwerkstatt, einem umfangreichen Konzept zur Förderung der Rechtschreibkompetenz, setzt die Gesamtschule seit dem Jahr 2010 ein spezielles Lernprogramm ein: *„'Graf Ortho' bringt schon den jüngsten Schülerinnen und Schülern den Umgang mit dem Computer und vor allem das sichere und zügige Schreiben im Zehnfingersystem näher”* (Gesamtschule Kamen 2011). In diesem Projekt erwerben die Mädchen und Jungen zusätzlich noch wichtige Informationen, wie sie den Computer als Werkzeug benutzen, um gefundene Informationen in eigenen Arbeiten zu verwenden und ihre Ergebnisse zu präsentieren. Dabei ist sich die Leitung der Gesamtschule stets bewusst, dass die Neuen Medien die traditionellen Medien nicht verdrängen sollen, sondern lediglich eine zeitgemäße Ergänzung darstellen können.

Für die Auswahl der Gesamtschule für das vorliegende Forschungsprojekt spricht zudem die Förderung mathematisch-naturwissenschaftlich interessierter Kinder bereits ab den Jahrgangsstufen 5 und 6. Diese werden zu „Naturwissenschafts-Profis” ausgebildet. Mit hohem Interesse besuchen sie zugleich Arbeitsgemeinschaften, in denen sie beispielsweise an Mathematikaufgaben knobeln, die mit der normalen Schulmathematik wenig zu tun haben. Im Mittelpunkt steht Kreativität und nicht die Durchführung von Standardoperationen, wie es im Unterricht zu Genüge getan wird. Wie beliebt diese AGs bei den Kindern sind, zeigt zum Beispiel der Zulauf, den die genannten Jahrgangsstufen beim „Känguru-Wettbewerb” 2013 verzeichneten. Ergänzt werden diese theoretischen Übungen durch praktische Besuche bei der Partnerfirma „Bayer” (vgl. ebd.).

5.4 Gesamtschule Süd, Essen

Die Gesamtschule Süd im Herzen des Ruhrgebiets in der 573.000-Einwohner-Metropole Essen hat vor allem damit zu kämpfen, dass *„Kinder aus südlichen Stadtteilen selten auf Hauptschulen und auch selten auf Gesamtschulen wechseln. Erhalten sie keine Empfehlung für ein Gymnasium, wechseln sie auf Realschulen"* (Stadt Essen 2011, S. 112). Im Unterschied zur Gesamtschule Kamen muss sich die Gesamtschule Süd daher eher mit einem hauptschulähnlichen Schülerklientel begnügen. Die Leiterin der Oberstufe, Ulrike Radar (2012), verweist zwar auch darauf, dass viele Kinder, die am Gymnasium nicht erfolgreich sind, irgendwann an der Gesamtschule Süd landen und dort sogar zu einem großen Teil ihr Abitur erreichen, doch bestreitet sie nicht, dass 45 % der Schülerschaft in ihrer Grundschulprognose als Hauptschülerinnen und -schüler abgestempelt wurden (zitiert nach Dowe 2012).

Da die nichtdeutschen Schülerinnen und Schüler in Essen zu mehr als einem Drittel zu Gesamtschulen wechseln, hat die Gesamtschule Süd zudem auch eine kulturell heterogene Schülerschaft zu betreuen. Schulleiter Menke (2012) bestätigt in einem Zeitungsinterview, dass *„hier viele Schüler aus bildungsfernen Schichten, die woanders abgelehnt wurden"* sind, doch *„Auch die haben ein Recht auf Beschulung"* (zitiert nach Dowe 2012).

Derzeit werden ungefähr 1.030 Schülerinnen und Schüler an der Gesamtschule Süd, die auch über eine gymnasiale Oberstufe und eine Ganztagsbetreuung verfügt, unterrichtet (vgl. Schulradar o.J.). Seit dem Frühjahr 2012 ist die Schule allerdings von der Schließung bedroht, da lediglich 34 Kinder von ihren Eltern angemeldet wurden. Somit konnte erstmals keine fünfte Jahrgangsstufe eingerichtet werden, da laut Schulgesetz 104 Neuanmeldungen nötig gewesen wären (vgl. Spletter 2012).

An der Gesamtschule Süd werden der fünfte und sechste Jahrgang zur Förderung und zur Orientierung genutzt. In diesen beiden Jahren findet der Unterricht vollständig im Klassenverband statt, so dass lediglich eine innere Differenzierung vorgenommen wird. Interessant ist allerdings die Differenzierung und Schwerpunktbildung im siebten und achten Jahrgang. Dies meint insbesondere, dass in den Fächern Mathematik und Englisch zwar die für Gesamtschulen übliche Fachleistungsdifferenzierung entsprechend des Leistungsstandes vorgenommen wird, doch werden die Kinder trotz der Aufteilung in Grund- und Erweiterungskurse weiterhin vorwiegend im Klassenverband unterrichtet (vgl. Schulprogramm Gesamtschule Süd). Das bedeutet, dass die ganze Klasse von den Lehrerinnen und Lehrern identisch unterrichtet wird, sich die zu bearbeitenden Aufgaben aber individuell an dem Leistungsvermögen der Kinder und Jugendlichen orientieren. Durch diese Art der Differenzierung erhofft sich die Gesamtschule Süd einen lockereren Mobilitätsprozess, vor allem hinsichtlich der Aufwärtsmobilität. Erst zu Beginn der neunten Jahrgangsstufe findet eine dem Leistungsniveau angepasste Klassenneubildung inklusive der Fachleistungsdifferenzierung in Deutsch und Chemie statt, um den unterschiedlichen Begabungen der Schülerschaft gerecht werden zu können. Ab diesem Zeitpunkt steht die Ausprägung von Abschlussprofilen im Fokus, so dass die Schülerinnen und Schüler entweder einen der möglichen Abschlüsse nach Klasse zehn erreichen oder aber die Oberstufe besuchen sollen (vgl. ebd.).

Hinsichtlich besonderer Mathematikangebote ist lediglich die gezielte Förderung leistungsschwacher Kinder im fünften und sechsten Jahrgang zu nennen. Die individuelle Förderung findet in Kleingruppen mit zwei Stunden pro Woche statt (vgl. ebd.).

Zudem wird jährlich im fünften Jahrgang eine Montessoriklasse eingerichtet, wobei im Rahmen fester Freiarbeitsstunden die Schülerinnen und Schüler mit entsprechenden Methoden (Wochenplanarbeit, frei bestimmtem Lerntempo, ...) alters- und entwicklungsgerecht gefördert werden (vgl. ebd.).

5.5 Gemeinschaftshauptschule an der Bruchstraße, Mülheim an der Ruhr

Während das Schulprogramm des Pestalozzi-Gymnasiums Unna beispielsweise die Schlagworte *„Wissen erweitern, Verantwortung übernehmen, Verständigung üben, Zukunft wagen"* verfolgt, sieht das an der städtischen Gemeinschaftshauptschule an der Bruchstraße, einer von vier Hauptschulen in der 167.000 Einwohner zählenden Ruhrgebietsstadt Mülheim an der Ruhr, deutlich anders aus. Während zwar auch die pädagogischen Leitbilder „Mitverantwortung" und „Vermittlung von Schlüsselqualifikationen" genannt werden, orientiert sich das Schulprogramm der Hauptschule aber viel eher an Grundsätzen, die ein Ausdruck von Menschlichkeit darstellen. Dabei handelt es sich beispielsweise um die „Identifikation mit der Schule", der „Stärkung des Selbstwertgefühls", der „Förderung der sozialen Kompetenz" oder der „Schaffung eines gewalt- und störungsfreien Arbeitsklimas" (Schulprogramm Hauptschule an der Bruchstraße).

Die Rahmenbedingungen an dieser Schule ermöglichen es nicht, den fachbezogenen Unterricht in den Mittelpunkt des täglichen Lebens zu rücken: Die Schule liegt im nördlichen Innenstadtbereich und wird derzeit von etwa 310 Schülerinnen und Schülern aus ganz Mülheim sowie Essen-Kettwig besucht. Von diesen 310 Kindern und Jugendlichen weisen 49 % (!) einen Migrationshintergrund aus 23 (!) verschiedenen Staaten auf. Dieser Wert liegt deutlich über dem landesweiten Durchschnitt an Hauptschulen in Nordrhein-Westfalen (vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen 2012) und führt dazu, dass sich auch die schulischen Förderprogramme eher auf sprachliche Disziplinen und die Integration der Immigrationskinder konzentrieren und weniger auf mathematische Lerninhalte abzielen.

Obwohl die 28 Lehrerinnen und Lehrer, unterstützt durch einen großen Teil der Schülerschaft, für den Erhalt und die Restaurierung des heruntergekommenen Schulgebäudes kämpfen, hat die Hauptschule an der Bruchstraße seit Jahren mit sinkenden Schülerzahlen und der drohenden Schließung zu kämpfen. Erst ein Bürgerentscheid konnte die Schule 2011 vor der eigentlich bereits beschlossenen Schließung retten.

Die Schulleiterin Gabriele Klar (2013) weist darauf hin, dass das Jahr 2013 zu einem Jahr großer Veränderungen werden sollte, in dessen Mittelpunkt die Umbenennung der Schule in „Max-Kölges-Schule am Dichterviertel - Partnerschule des Handwerks" stand. Max Kölges hatte starken Einfluss auf die Prägung des Mülheimer Handwerks und gilt damit als ideale Identifikationsfigur für die berufsorientierte Ausrichtung der Hauptschule. Ziel der Initiative ist es, jeder Schülerin und jedem Schüler einen Ausbildungsplatz zu verschaffen.

Betrachtet man die Entwicklungsziele für das Kernfach Mathematik, versucht auch die Hauptschule an der Bruchstraße die Schülerinnen und Schüler an sogenannte „Schlüsselaufgaben", also komplexe Aufgaben mit hohem Wirklichkeitsbezug, zu gewöhnen. Darüber hinaus soll eine größere Nachhaltigkeit des mathematischen Wissens dadurch erreicht werden, indem der Lernstoff verschiedener Jahrgänge flexibler in das Unterrichtsgeschehen integriert wird. In diesem Sinn ist auch eine am Ende jedes Schuljahres geschriebene, bezüglich der Themen, jahrgangsübergreifende Klassenarbeit zu verstehen, die den Schülerinnen und Schülern zu mehr Sicherheit im Umgang mit mathematischen Themen, auch im späteren Berufsleben, verhelfen soll (vgl. Schulprogramm Hauptschule an der Bruchstraße).

Hinsichtlich des Umgangs mit Neuen Medien hat die Hauptschule an der Bruchstraße ebenfalls ein eigenes Konzept entwickelt, da gerade von Hauptschülerinnen und Hauptschülern beim Berufseinstieg zumindest ein sicherer Umgang mit dem Computer erwartet wird. Das Medienkonzept soll den Schülerinnen und Schülern dabei sowohl die Welt der Neuen Medien näherbringen und einen zielgerichteten Einsatz ermöglichen als auch mit Hilfe entsprechender Lernprogramme und Lernspiele den Fachunterricht vorantreiben (vgl. ebd.).

5.6 Ein zusammenfassender Überblick über die Stichprobe

Das Pestalozzi-Gymnasium Unna ist sicherlich die anspruchsvollste Schule, welche an dem Forschungsprojekt teilnimmt. Gerade dadurch ist ein informativer und spannender Kontrast zu den Forschungsergebnissen der Gesamt- und Hauptschulen zu erwarten. Aufgrund der ausgewogenen sozialen Zusammensetzung, einer überdurchschnittlichen Abiturnote und einem offenen Umgang mit den Neuen Medien eignet sich die Gesamtschule Kamen, die sich eher auf Realschulniveau bewegt, für die Erhebung der Daten bezüglich des Fernsehverhaltens und dem untersuchten Einfluss auf die Rechenfähigkeit ebenfalls in besonderem Maße. Im Gegensatz dazu hat die Gesamtschule Süd mit den beschriebenen Problemen zu kämpfen, so dass die Schule zu einem interessanten Gegenpol innerhalb der beiden Gesamtschulen wird und das Spektrum damit noch um eine weitere Komponente ergänzt. Es ist deutlich geworden, dass die Hauptschule an der Bruchstraße ganz andere Sorgen hat, als Schülerinnen und Schüler auf Mathematikwettbewerbe vorzubereiten. Ganz im Gegenteil, hier geht es einzig und allein darum, die Jugendlichen dazu zu befähigen, einen Ausbildungsplatz zu erhalten. Demnach erweitert die Hauptschule die Bandbreite um eine nötige weitere Dimension.

Die 232 Schülerinnen und Schüler, die an dem Forschungsprojekt teilnehmen, verteilen sich folgendermaßen auf die vier Schulen:

Tabelle 5.1

Zusammensetzung der Probanden

SchülerInnen	Pestalozzi-Gymnasium	Gesamtschule Kamen	Gesamtschule Süd	Hauptschule a. d. Bruchstr.
absolute Häufigkeit	90	57	47	38
		Gesamtschule: 104		
in Prozent	38,7	44,8		16
in NRW in Prozent	42	19,4 (inkl. Realschule: 43,3)		14,6

Tabelle 5.1 gibt dabei zunächst die absoluten Häufigkeiten der Probanden pro Schulform wieder. Insgesamt nehmen 90 Schülerinnen und Schüler aus allen drei siebten Jahrgängen des Pestalozzi-Gymnasiums, 104 Schülerinnen und Schüler aus vier Gesamtschulklassen und 38 Schülerinnen und Schüler aus zwei Klassen der Hauptschule an der Bruchstraße an der Studie teil. Wesentlich interessanter und aussagekräftiger sind aber die nächsten beiden Zeilen. Demnach sind 38,7 % der an der Studie teilnehmenden Jugendlichen Gymnasiasten, 44,8 % GesamtschülerInnen und 16 % HauptschülerInnen. Um die Repräsentativität der vorliegenden Probandenverteilung zu überprüfen, sind zudem die relativen Häufigkeiten aus ganz Nordrhein-Westfalen für die verschiedenen Schulformen angegeben. Direkt ablesbar ist, dass die Häufigkeiten für die Hauptschule und das Gymnasium ziemlich passend in dem Forschungsprojekt wiedergegeben werden. Nur die Gesamtschule scheint völlig überrepräsentiert zu sein. Zieht man nun aber die Annahme hinzu, dass die Gesamtschule Kamen dem Real-

schulniveau relativ nahekommt, es sich also gemessen an den statistischen Werten nun eher um eine Real- als um eine Gesamtschule handelt, entspricht die relative Häufigkeit ziemlich exakt der Anzahl der Schülerinnen und Schüler die in ganz NRW entweder die Gesamt- oder aber die Realschule besuchen.

Somit ist zusammenfassend erkennbar, dass die prozentuale Verteilung der Probanden auf die verschiedenen Schulformen die prozentuale Verteilung aller nordrhein-westfälischen Schülerinnen und Schüler auf die entsprechenden Schulformen überraschend genau abbildet, so dass die Stichprobe bzw. die Grundgesamtheit durchaus eine gewisse Repräsentativität besitzt.

Zudem besteht die Stichprobe zu 54,3 % aus männlichen und zu 45,7 % aus weiblichen Probanden.

Abschließend soll kurz auf ein weiteres wichtiges Detail eingegangen werden. Nicht zu unterschätzen ist bei der Auswahl der Stichprobe nämlich der Umgang mit dem Datenschutz. Zwar findet dieses heikle Thema in der Literatur nur selten die nötige Aufmerksamkeit, doch sind bereits einige interessante Forschungsprojekte an der Datenschutzfrage gescheitert (vgl. KIRCHHOFF et al. 2006, S. 17).

Auch für das vorliegende Projekt ist der Umgang mit dieser Problematik von Bedeutung, da sämtliche Schulleitungen auf die Sicherstellungen des Datenschutzes insistierten. Schnell stellt sich die Frage, wie eine anonyme Datenerhebung der Fernsehgewohnheiten den konkreten Rechenfähigkeiten der Jugendlichen zugeordnet werden kann, ohne den Datenschutz zu verletzen. Aus diesem Grund soll kurz auf das Vorgehen eingegangen werden: Die Prozedur basiert vor allem darauf, dass jeder befragten Schülerin und jedem befragten Schüler eine feste Identitätsnummer für den kompletten Verlauf des Projekts zugeordnet wird. Nur dem jeweiligen Lehrkörper ist bekannt, welche Nummer zu welchem Probanden gehört, so dass die Leistungen in Mathematik anonymisiert über die Identitäts-Nummern angegeben und dem Verfasser übermittelt werden können. Da die Lehrkräfte im Gegenzug nicht über die konkreten Antworten ihrer Schülerinnen und Schüler auf die Fragen in den Fragebögen in Kenntnis gesetzt werden, findet eine strikte Trennung zwischen den Informationen, welche den Lehrerinnen und Lehrern einerseits und dem Verfasser andererseits zugänglich gemacht werden statt. Durch diese Anonymisierung erwartet sich der Verfasser ein möglichst wahrheitsgemäßes Ausfüllen der Fragebögen.

6. Datenerhebung und erste Ergebnisse

6.1 Ein Überblick über die zu erhebenden Daten

Nachdem bisher die Frage bezüglich der Fallauswahl, also derjenigen Probanden, über die Aussagen getroffen werden sollen, zufriedenstellend beantwortet werden konnte, soll nun begründet werden, was und in welcher Form gefragt bzw. erfasst wird.

Um die zentrale Fragestellung, ob es einen Zusammenhang zwischen den Mathematikleistungen und dem Fernsehverhalten gibt, zu untersuchen, ist die Entscheidung getroffen worden, fragebogenbasiert vorzugehen und damit auf Einzelinterviews der Probanden zu verzichten, auch weil sich insbesondere aufgrund der Stichprobengröße von deutlich über 200 Schülerinnen und Schülern umfangreiche und detaillierte Interviews für eine einzige Testperson nahezu unmöglich in ihrer Durchführung gestalten. Formuliert man es zudem leicht überspitzt, kommt man einerseits ohnehin zu dem Ergebnis, dass die zu benutzenden Fragebögen letztlich *standardisierten Interviews* sehr nahe kommen sowie andererseits, dass selbst die mathematischen und methodischen Schwierigkeiten, die eine Fragebogenaktion mit nachfolgender datenanalytischer Auswertung mit sich bringt, durch Einzelinterviews mitnichten vermieden würden. Die Unterschiede zwischen den beiden alternativen Vorgehensweisen liegen dementsprechend nur in der Datenerhebung, nicht aber im Ergebnis und der Auswertung. Da nun aber durch die Fragebögen eine wesentlich größere Probandengruppe untersucht werden kann, eignet sich dieses Verfahren in besonderem Maße und soll fortan Verwendung finden.

Wenngleich ein Fragebogen zunächst als eine recht alltägliche Angelegenheit erscheint, müssen dennoch zahlreiche logische Regeln bei der Konstruktion eines Fragebogens beachtet werden. Schließlich müssen die unterschiedlichen Fragen aufeinander abgestimmt, auf die Bedürfnisse und Fähigkeiten der Zielgruppe maßgeschneidert und in Einklang mit den Forschungszielen gebracht werden (vgl. Kirchhoff et al. 2006, S. 19). Bei der Erstellung des Fragebogens verweist die Literatur zusätzlich darauf, dass in der Fragebogenkürze die Ausfüllwürze liege und man sich aus diesem Grund kurz fassen solle (vgl. ebd., S. 27). Schließlich müssen die Fragebögen von jedem Probanden ohne Zeitdruck innerhalb der festgelegten Zeitspanne vollständig auszufüllen sein. Um all diesen Aspekten gerecht zu werden, muss daher zunächst die durchdachte Entscheidung getroffen werden, welche Variablen erhoben werden sollen.

In Erinnerung an die im dritten Kapitel vorgestellten empirischen Studien, die sich in das Themenfeld „Fernsehverhalten und Bildung" einsortieren lassen, fällt auf, dass die Wissenschaftlerinnen und Wissenschaftler meist als Variable entweder die Fernsehdauer oder aber die Art der geschauten Fernsehsendungen untersuchen. PETER WINTERSTEINS *Mensch-Zeichen-Test* von 2004 erforscht zum Beispiel den Einfluss von exzessivem Fernsehkonsum auf die bildhafte Wahrnehmung von Kindern (vgl. WINTERSTEIN; JUNGWIRTH 2006), während das *Sesamstraßenexperiment* die Effekte einer explizit für das Forschungsprojekt entwickelten und somit pädagogisch geprägten Fernsehsendung auf die Lese-, Schreib- und Rechenfähigkeiten von Kindern wiedergibt (vgl. LEE 2009). Die Ergebnisse deuten darauf hin, dass zeitlich überhöhter Fernsehkonsum für Kinder und Jugendliche schädlich sein kann, aber inhaltlich anspruchsvolles Fernsehen die kindliche Entwicklung sogar zu fördern vermag, so dass sowohl die Dauer als auch der Inhalt sich tatsächlich als wirkmächtige Faktoren herauskristallisiert haben. Leider spielen in den wenigsten der bisherigen Studien aber beide Variablen eine gleichbedeutende Rolle. Um eben dieses Vakuum in dieser Arbeit weitestgehend zu schließen, werden folgerichtig sowohl die Dauer als auch die Art des Fernsehkonsums Forschungsgegenstand sein. Weiterhin erscheint es interessant, ob die sonstige Freizeitgestaltung der Probanden ebenfalls einen Einfluss auf die Leistungen in Mathematik hat. Schließlich ist in der von MANFRED SPITZER (2012) zitierten Studie von MYRTEK (2003) eine sehr genaue Beschreibung des Alltags der Probanden vorgenommen worden, welche neben den Fernsehgewohnheiten eben auch die übrigen Freizeitbeschäftigungen berücksichtigte. Wie beschrieben, stellte sich dabei heraus, dass diejenigen Probanden, die viel fernsehen, sich weniger bewegen, weniger lesen, häufiger allein sind und seltener Gespräche. Basierend auf diesen Ergebnissen lässt sich ein zusätzlicher Einfluss der Freizeitgestaltung sowohl auf die Mathematikleistungen als auch auf die Fernsehgewohnheiten der Schülerinnen und Schüler eben nicht per se ausschließen, so dass auch die fernsehfreien Freizeitaktivitäten einen Platz in der empirischen Studie des Verfassers finden sollen. Damit umfasst die Umfrage zum Fernsehverhalten die drei Themenschwerpunkte

- Dauer bzw. Häufigkeit des Fernsehkonsums,
- Art bzw. latente Faktoren des Fernsehkonsums,
- sonstige Freizeitaktivitäten,

die getrennt voneinander abgefragt werden. Der vierte Erhebungspunkt ist letztlich natürlich die Mathematikleistung der Probanden und wird im nachfolgenden Kapitel methodisch aufwendig präsentiert.

6.2 Dauer und Häufigkeit des Fernsehkonsums

Sozusagen zum Aufwärmen, den Befragten sollte die Chronologie eines Fragebogens vertraut sein, um das Ausfüllen zu erleichtern (vgl. Kirchhoff et al. 2006, S. 19), beginnt die Umfrage jedoch mit Fragen zum Alter, der Klassenzugehörigkeit und dem Geschlecht der Kinder. Da in der geschlechtsspezifischen Medienkompetenzforschung zahlreiche Unterschiede zwischen Jungen und Mädchen in der Nutzung digitaler Medien entdeckt wurden, erscheint die Unterscheidung der Fragebögen für die kommende Auswertung durchaus interessante Ergebnisse zu versprechen (vgl. Luca; Aufenanger 2007, S. 45).

Den Schülerinnen und Schülern werden im Anschluss zwei Fragen hinsichtlich ihrer zeitlichen Fernsehnutzgewohnheiten gestellt. Selbstverständlich hängt der geplante Aufbau der Fragen in erster Linie von inhaltlichen Überlegungen ab. Das heißt allerdings nicht, dass optische Reize und Aspekte keine Rolle spielen. Einerseits erleichtert eine ansprechende und in diesem Fall jugendgerechte Aufmachung das Erfassen bestimmter Fragen. Andererseits verdeutlicht die optische Gestaltung den Probanden, ob sie Mühen wert sind und als Subjekt wahr- und ernst genommen oder lediglich als Objekt einer Untersuchung betrachtet werden (vgl. Kirchhoff et al. 2006, S. 25). Aus diesem Grund hat der Verfasser entschieden, an passenden Stellen comicartige *Cliparts* in den Fragebogen zu integrieren und sämtliche Fragebögen farbig drucken zu lassen, um die Kinder zu motivieren, mit Eifer und Ehrlichkeit an dem Forschungsprojekt teilzunehmen. Um dies zu verdeutlichen, werden die entsprechenden *Cliparts* auch in diese Arbeit eingebunden.

6.2.1 Datenerhebung

Auf den ersten Blick erscheint es zwar relativ simpel, die Fernsehdauer sowie die Fernsehhäufigkeit der Probanden zu bestimmen, doch verweist die Literatur gerade bei vordergründig banalen Fragen zu Verhaltenshäufigkeiten darauf, besonders achtsam zu sein. Aus diesem Grund gilt es zunächst explizit zu durchdenken, welcher Fragentyp genutzt werden soll, um ein optimales Ergebnis zu erzielen. Grundsätzlich werden Fragen mit gebundenen und freien Antworten unterschieden. Dabei geben gebundene Fragestellungen konkrete Antwortalternativen vor, während bei freien Fragestellungen eben keine bzw. nur wenige Einschränkungen existieren (vgl. Bühner 2006, S. 45). Beide Antworttypen haben sowohl Vor- als auch Nachteile.

Die erste Frage verfolgt nun das Ziel, herauszufinden, wie viele Stunden die Jugendlichen am Tag vor dem Fernseher verbringen. Dabei fällt die Wahl, wie unten abgebildet, auf eine freie, aber teilkonstruierte Fragestellung in Form einer Ergänzungsaufgabe. Die Probanden müssen selbstständig durch Eintragen einer Stundenzahl in das dafür vorgesehene freie Feld den Satz *„Am Tag schaue ich ungefähr ◯ Stunden Fernsehen!"* ergänzen. Becker, Spörrle und

FÖRSTERLING (2003, S. 46) sprechen sich nämlich dafür aus, bei Fragen zu Verhaltenshäufigkeiten konkret nach Quantitäten zu fragen und keine Antwortvorgaben zu machen. Beachtet werden muss dabei allerdings, dass die Schülerinnen und Schülern bei derartigen Fragen möglicherweise ungenaue und vor allem subjektive Antworten geben könnten. Die Genauigkeit der Beantwortung hängt vor allem von der Motivation der Probanden ab. Um die Kinder zu motivieren, bietet sich die verwendete Grafik an dieser Stelle besonders gut an.

Abbildung 6.1 Ausschnitt aus dem Fragebogen (1)

Mit der zweiten Frage soll untersucht werden, an welchen Tagen die Kinder sich vorzugsweise mit dem Fernsehen beschäftigen. Diese Frage ist durchaus hilfreich, um zu erkennen, ob ein Jugendlicher, welcher bei der ersten Frage beispielsweise eine hohe Stundenzahl angegeben hat, täglich soviel fernsieht oder möglicherweise nur an ausgewählten Tagen, wie beispielsweise dem Wochenende, wodurch sich die gegebene Antwort aus Frage 1 relativieren würde. Entsprechend ist aber auch das Gegenteil möglich. Ein Proband, der bei Frage 1 einen relativ geringen Wert angegeben hat, aber jeden Tag in der Woche vor dem Fernseher sitzt, schaut womöglich in der Woche viel länger fern als andere Probanden, die zwar lange, aber dafür nur selten schauen. Für diese Frage eignet sich natürlich nur eine gebundene Fragestellung mit den sieben Wochentagen als vorgegebene Antworten. Die Schülerinnen und Schüler müssen lediglich ankreuzen, an welchen Tagen sie normalerweise bzw. bevorzugt fernsehen.

Frage 2: An welchen Tagen schaust du Fernsehen? Kreuze an!

Montag	Dienstag	Mittwoch	Donnerstag	Freitag	Samstag	Sonntag

Abbildung 6.2 Ausschnitt aus dem Fragebogen (2)

Basierend auf diesen beiden Fragen lässt sich nun leicht die durchschnittliche Dauer des täglichen Fernsehkonsums jedes Probanden mit Hilfe der Formel $\frac{Stundenzahl \cdot Tage}{7}$ ermitteln.

Nun mag es an dieser Stelle durchaus nachvollziehbare Einwände gegen die Genauigkeit dieser Form der Datenerhebung geben. Neben der Tatsache, dass die JIM-Studie auf vergleichbare Art und Weise seit vielen Jahren zuverlässiges Datenmaterial gewinnt, legitimiert auch die geplante Auswertungsmethode dieser Studie ein entsprechendes Vorgehen. Um die Abweichung der durchschnittlichen Dauer des täglichen Fernsehkonsums der befragten Jugendlichen von der durchschnittlichen Dauer des täglichen Fernsehkonsums ihrer Altersgruppe zu ermitteln, wird eine Clusteranalyse, mit dem Ziel die Probanden in optimale Gruppen ähnlicher Fernsehnutzungsdauer einzuteilen, zum Einsatz kommen. Es ist also das Ziel des Verfassers, zu ermitteln, welche Probanden eher wenig, mittel oder viel Zeit vor dem Fernseher verbringen. Daher ist eine vollkommen exakte Zeitangabe überhaupt nicht notwendig, schließlich kann davon ausgegangen werden, dass vor allem die Extremgruppen hinreichend genau eingeteilt werden können. Der Verfasser ist davon überzeugt, dass Probanden, welche außerordentlich viel bzw. besonders wenig fernsehen, dies mit ihren Zeitangaben zum Ausdruck bringen werden.

6.2.2 Wenigseher, Mittelseher, Vielseher, Exzessivseher

Nachdem nun die Daten erhoben und in die durchschnittliche tägliche Fernsehdauer transformiert worden sind, gilt es fortan auf mathematisch korrekte Art und Weise die verschiedenen Gruppen auf Grundlage des Datensatzes zu generieren. Ein häufiges statistisches Auswertungsverfahren um Merkmale zu ähnlichen Gruppen zu fusionieren sind Methoden der Clusteranalyse, welche im vierten Kapitel kurz vorgestellt wurden. Aufgrund seiner positiven Eigenschaften bezüglich der Analyse stetiger Datensätze soll zur Gruppierung der Probanden der Ward-Algorithmus als Clustertechnik verwendet werden (vgl. ebd., S. 285). Als Proximitätsmaß erfordert der Ward-Algorithmus beispielsweise die quadrierte Euklidische Distanz, die bekanntermaßen durch die Formel $d^2(x,y) = \sum_{i=1}^{n}(x_i - y_i)^2$ berechnet wird. Der Ward-Algorithmus gehört zu den konservativen Clusterverfahren.

Üblicherweise laufen die hierarchischen Verfahren der Clusteranalyse recht ähnlich ab. Ausgehend von n Clustern wird im ersten Schritt, basierend auf dem Proximitätsmaß eine Distanzmatrix berechnet. Im Anschluss werden dann die beiden Cluster vereinigt, die im Sinne des gewählten Verfahrens am ähnlichsten sind. Im Anschluss wird auf Grundlage des neuen Datensatzes mit $n-1$ Clustern eine neue Distanzmatrix berechnet und der Vorgang beginnt von vorne.

Die Ward-Methode basiert nun darauf, dass jene Cluster fusioniert werden, die zu einem minimalen Homogenitätsverlust in Bezug auf das Varianzkriterium, definiert als die Summe der quadratischen Euklidischen Abstände der Gruppenelemente zu ihren Gruppenmittelwerten, die über alle Gruppen berechnet wird, führen. Häufig wird das Varianzkriterium auch als Fehlerquadratsumme bezeichnet. Es ist als Zuwachsgröße beim Ward-Verfahren zwar nicht global, aber für jede Fusionsstufe als vorgegebene Zielfunktion des Klassifizierungsprozesses zu minimieren. Dementsprechend werden die beiden Cluster A_r und A_s zusammengefasst, für die der Ausdruck $\frac{n_r \cdot n_s}{n_r + n_s} || \overline{x}_r - \overline{x}_s ||^2$ minimiert wird, so dass die Varianz innerhalb des neuen Clusters A_g möglichst langsam wächst (vgl. Müller 2010, S. 137).

Zwar wird in der Literatur davor gewarnt, dass die Anwendung eines hierarchischen Fusionierungsalgorithmus, der lediglich schrittweise Optimierung impliziert, nicht automatisch optimale Partitionen im Sinne des Varianzkriteriums zu bilden vermag (vgl. Hudec 2003, S. 3-7), da sich die Optimalität bezüglich des Varianzkriteriums nur auf die Zusammenfassung bezüglich der einzelnen Iterationsschritte bezieht (vgl. Uni Karlsruhe (o.V.) o.J.), doch zeigte sich in einem Simulationsexperiment zur Strukturanalyse von Bergs (1980, S. 96 zitiert nach Tauberger 2008), dass mittels des Ward-Verfahrens Elemente mit einer Genauigkeit von 97,12 % der *richtigen* Gruppe zugeordnet werden konnten. Zusammenfassend ergab das Experiment, dass der Ward-Algorithmus sehr gute Cluster und meistens sogar die für das Experiment angemessene Clusterzahl finden konnte.

Darüber hinaus ist beim Ward-Verfahren zu beachten, dass Ausreißer im Datensatz nicht erkannt werden, da sie meist recht früh einem endgültigen Cluster zugewiesen werden (vgl. Uni Karlsruhe (o.V.) o.J.), so dass es sich empfiehlt, den Datensatz zum Beispiel mit Hilfe der Regel von Tukey auf Ausreißer zu untersuchen. Bei dieser Regel werden diejenigen Daten x_i mit $i \in \{1,...,n\}$ einer Datenreihe $x_1,...,x_n$ mit den dazugehörigen Quantilen q_1, q_3 und dem Inter-Quartile-Range $iqr(x_1,...,x_n) = q_3 - q_1$ als Ausreißer bezeichnet, für die $x_i < q_1 - 1,5iqr$ oder $x_i > q_3 + 1,5iqr$ gilt.

Nachdem nun die Entscheidung gefallen ist, den Ward-Algorithmus als Fusionierungsmethode zu benutzen, kann die Auswertung durchgeführt werden. Wie bei jeder statistischen Methode, die von einem Anwender genutzt wird, sollte allerdings der Datensatz nicht aus dem Blick geraten, so dass zunächst einige Beobachtungen festgehalten werden sollen.

	Schulform	Gesamtschule											
	Proband	P001	P002	P003	P004	P005	P006	P007	P008	P009	P010	P011	P012
	Geschlecht	m	m	m	m	m	m	w	w	w	m	m	m
Frage 1:	Dauer	3	2,25	3,75	2,25	2,25	2,25	1,88	3,38	5,25	2,25	1,88	1,5
Frage 2:	Häufigkeit	7	7	6	6	5	4	7	7	7	6	7	4
	tgl. Dauer	3	2,25	3,21	1,93	1,61	1,29	1,88	3,38	5,25	1,93	1,88	0,86

Abbildung 6.3 Ausschnitt aus dem Datensatz

Abbildung 6.3 stellt einen Auszug aus dem Datensatz dar. Es ist zu erkennen, dass die Schulform, die Identitätsnummer, das Geschlecht, das Beantwortungsergebnis der beiden Fragen sowie die berechnete tägliche Fernsehdauer jedes Probanden festgehalten sind. Insgesamt haben die Probanden angegeben, dass sie durchschnittlich pro Tag **2 Stunden 13 Minuten** fernsehen. Die Standardabweichung liegt bei **1,43**. Vertraut man lieber dem gegen Ausreißer weniger anfälligen Median, schauen die Schülerinnen und Schüler täglich **2 Stunden 15 Minuten** fern. Beide Werte scheinen ziemlich zufriedenstellend zu sein. Die JIM-Studie von 2011 gibt beispielsweise einen Durchschnittskonsum von **1 Stunde 51 Minuten** an, bezieht aber das Wochenende nicht mit ein (vgl. mpfs 2013b, S. 25).

Detailliertere Informationen liefert die nachfolgende Tabelle:

Tabelle 6.1

Durchschnittliche Fernsehdauer nach verschiedenen Kategorien

Insgesamt	Gymnasium	Gesamtschule	Hauptschule	männlich	weiblich
2h 13min	1h 29min	2h 32min	3h 2min	2h 12min	2h 14min

Tabelle 6.1 präsentiert jeweils die durchschnittliche Fernsehdauer, unterteilt in die Kategorien *Schulform* und *Geschlecht*.

Demnach schauen die Gymnasiasten durchschnittlich am wenigsten fern (**1 Stunde 29 Minuten**), während die Probanden der Gesamtschule bereits **2 Stunden 32 Minuten** fernsehen. Übertroffen werden diese aber sogar noch von den Schülerinnen und Schülern der Hauptschule, die täglich im Schnitt knapp über **3 Stunden** vor dem Fernsehgerät sitzen. Sowohl die Gesamt- als auch die HauptschülerInnen liegen damit oberhalb des Durchschnittswerts aller Probanden. Auch diese Feststellungen stimmen mit den Ergebnissen der letzten JIM-Studie überraschend präzise überein. Dort wurde beispielsweise für Gymnasiasten eine Fernsehdauer von **1 Stunde 36 Minuten** erfasst (vgl. ebd., S. 25).

Zwischen Mädchen und Jungen besteht kaum ein Unterschied. Beide Geschlechter schauen in etwa gleich viel fern und liegen dementsprechend eng am Durchschnitt der gesamten Stichprobe.

Bevor nun der Ward-Algorithmus angewendet werden kann, muss der Datensatz abschließend noch auf etwaige Ausreißer untersucht werden, da die Ward-Methode, wie erklärt, sehr anfällig gegenüber Ausreißern ist. Hilfreich ist dabei ein Boxplot-Diagramm.

Ein Boxplot gilt als stark vereinfachte, aber dafür umso informativere Darstellung eines Datensatzes. Die Illustration basiert auf den Quantilen und macht eben auch Ausreißer deutlich. Die Darstellung wurde 1977 von John W. Tukey entwickelt. Dabei liegt die untere Kante des Vierecks, also der „Box", auf der Höhe des 25-%-Quantils und die obere Kante auf der Höhe des 75-%-Quantils. Der Querstrich innerhalb des Rechtecks gibt den Median an. Die beiden Stacheln, die von der unteren und der oberen Kante ausgehen, verbinden diese Kanten mit der letzten Beobachtung, die nach der oben genannten Ausreißerregel von Tukey kein Ausreißer ist. Die Kringel ober- oder unterhalb dieser Stacheln gelten als Ausreißer (vgl. Meise o.J., S. 33). Abbildung 6.4 zeigt das zu dem Datensatz gehörende Boxplot:

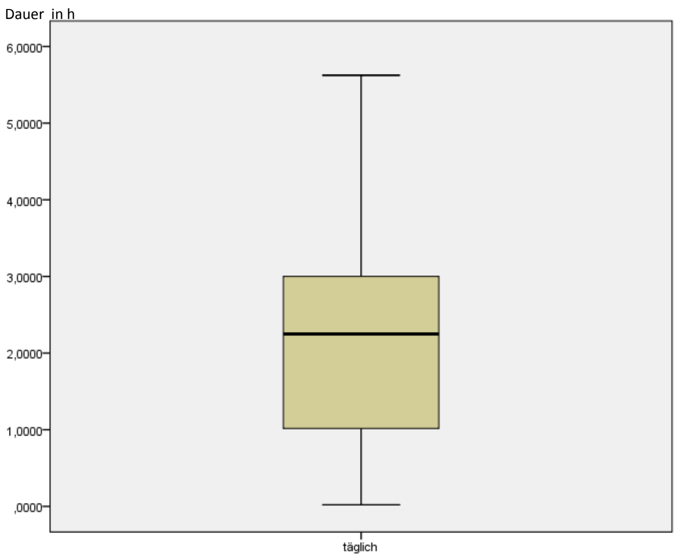


Abbildung 6.4 Boxplot der Fernsehdauer

Der Median liegt in dem vorliegenden Fall bei **2 Stunden 15 Minuten**. In der Box liegen die Daten, die durch das Intervall von **einer** bis **drei Stunden** beschrieben werden. Da allerdings in der Grafik keine Kringel zu finden sind, kann geschlossen werden, dass keine Ausreißer vorliegen. Dementsprechend symbolisiert das untere Ende bzw. das obere Ende des Boxplots zugleich das Minimum bzw. Maximum der Datenreihe. Die geringste genannte tägliche Fernsehdauer beträgt demnach **1 Minute**, während der Maximalwert bei **5 Stunden 38 Minuten** liegt.

Dieser erste Überblick lässt auf eine erfolgreiche und positive Datenerhebung schließen, so dass der Clusteranalyse nach Ward zur Ermittlung der jeweiligen Gruppen ähnlicher Fernsehnutzungsdauer nichts mehr im Wege steht. Das Verfahren wird in SPSS durchgeführt und durchläuft die nachfolgenden Schritte.

Zusammenfassung der Fallverarbeitung[a]

Fälle					
Gültig		Fehlend		Gesamt	
N	Prozent	N	Prozent	N	Prozent
232	100,0	0	,0	232	100,0

a. Ward-Linkage

Abbildung 6.5 Zusammenfassung der Fallverarbeitung

Die Tabelle „Zusammenfassung der Fallverarbeitung” gibt zunächst die Anzahl der gültigen und fehlenden Daten an. Da alle Probanden eine Zeitangabe auf ihren Fragebögen hinterlassen haben, existieren keine fehlenden Werte, so dass SPSS diese nicht simulieren muss. Die Clusteranalyse wird also nach dem Ward-Verfahren durchgeführt und umfasst die 232 Antworten der Probanden bei null fehlenden Werten.

Näherungsmatrix

Quadriertes euklidisches Distanzmaß												
110	111	112	113	114	115	116	117	118	119	120	121	122
,141	2,250	3,516	,000	1,389	6,612	6,891	3,318	,000	7,760	8,369	,563	2,25
,141	,563	6,891	,563	3,719	3,318	11,391	6,612	,563	4,144	4,592	2,250	5,06
,347	2,939	2,758	,046	,930	7,760	5,812	2,583	,046	9,000	9,654	,287	1,65
,485	,184	8,681	1,148	5,063	2,250	13,664	8,369	1,148	2,939	3,318	3,318	6,61
1,036	,011	10,679	1,940	6,612	1,389	16,143	10,332	1,940	1,940	2,250	4,592	8,36
1,794	,046	12,883	2,939	8,369	,735	18,829	12,501	2,939	1,148	1,389	6,073	10,33
,563	,141	9,000	1,266	5,306	2,092	14,063	8,681	1,266	2,758	3,125	3,516	6,89
,563	3,516	2,250	,141	,646	8,681	5,063	2,092	,141	9,990	10,679	,141	1,26
6,891	14,063	,141	5,063	1,148	23,246	,141	,184	5,063	25,358	26,449	2,250	,56
,485	,184	8,681	1,148	5,063	2,250	13,664	8,369	1,148	2,939	3,318	3,318	6,61
,563	,141	9,000	1,266	5,306	2,092	14,063	8,681	1,266	2,758	3,125	3,516	6,89
3,125	,413	16,143	4,592	11,032	,184	22,732	15,716	4,592	,413	,563	8,369	13,27

Abbildung 6.6 Näherungsmatrix

Die hier nur auszugsweise abgebildete Tabelle „Näherungsmatrix” enthält für die geclusterten Fälle die paarweisen Abstände in Form einer Matrix, basierend auf dem gewählten Distanzmaß, so dass sich in der Tabelle alle nach Euklid berechneten Abstände zwischen sämtlichen Probanden befinden. Zu erwähnen ist noch, dass es sich hierbei um eine Unähnlichkeitsmatrix handelt. Hohe Werte drücken eine große Unähnlichkeit und damit eine geringe Ähnlichkeit aus (vgl. SCHENDERA 2010, S. 56).

Ward-Linkage

Zuordnungsübersicht

Schritt	Zusammengeführte Cluster		Koeffizienten	Erstes Vorkommen des Clusters		Nächster Schritt
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	175	232	,000	0	0	194
2	209	231	,000	0	0	24
3	221	230	,000	0	0	12
4	223	229	,000	0	0	10
5	217	228	,000	0	0	16
6	149	227	,000	0	0	77
7	185	226	,000	0	0	44
8	215	225	,000	0	0	18
9	173	224	,000	0	0	55
10	5	223	,000	0	4	56
11	220	222	,000	0	0	13
12	39	221	,000	0	3	28
13	12	220	,000	0	11	15
14	218	219	,000	0	0	15
15	12	218	,000	13	14	81

Abbildung 6.7 Zuordnungsübersicht (Anfang)

In der ebenfalls nur auszugsweise abgebildeten „Zuordnungsübersicht" sind die einzelnen Schritte der Clusteranalyse dargestellt. Die Übersicht zeigt damit den Verlauf der Clusterbildung von der ersten (alle Objekte sind isoliert) bis zur letzten Stufe (alle Objekte sind in einem Cluster vereint) (vgl. KOPP; LOIS 2009, S. 47). In der Spalte „Schritt" sind die $k-1$ ($N = 231$) Schritte aufgelistet, in deren Verlauf die k Probanden ($N = 232$) zusammengeführt werden. Die Spalte „Zusammengeführte Cluster" führt nacheinander auf, welche beiden Cluster zu welchem Zeitpunkt zusammengeführt werden. In Schritt 1 werden dementsprechend Proband 175 und Proband 232 zu einem Cluster fusioniert. Dementsprechend kann davon ausgegangen werden, dass diese beiden Probanden eine sehr ähnliche tägliche Fernsehdauer angegeben haben. Nach demselben Muster werden in Schritt zwei die Probanden 209 und 231 zusammengeführt. Dieses Verfahren wiederholt sich bis zum Ende (vgl. SCHENDERA 2010, S. 56).

Die Spalte „Koeffizienten" gibt die zunehmende Heterogenität an (vgl. ebd., S. 56). Sie wächst stetig, da zunächst die Cluster mit der größten Ähnlichkeit verschmolzen werden, dann diejenigen mit der zweitgrößten usw. Die Distanzwerte steigen zunächst äußerst moderat an. Erst auf den späteren Stufen ist ein deutlicher Anstieg zu beobachten. In vorliegendem Fall nimmt die Heterogenität erst ab dem 191. Schritt spür- und messbar zu. Dieser Verlauf ist nicht untypisch und lässt sich leicht erklären: Auf den unteren Stufen werden noch viele Cluster mit hohen Ähnlichkeiten gebildet, während im weiteren Verlauf auch ursprünglich unähnliche Cluster zusammengeführt werden. Besonders deutlich wird dies ab Schritt 229 (vgl. KOPP; LOIS 2009, S. 49).

223	2	4	5,589	184	211	227
224	1	14	8,041	221	187	229
225	56	67	11,684	205	220	228
226	12	49	16,084	222	218	230
227	2	5	22,195	223	214	229
228	9	56	33,365	219	225	231
229	1	2	62,968	224	227	230
230	1	12	181,596	229	226	231
231	1	9	472,085	230	228	0

Abbildung 6.8 Zuordnungsübersicht (Ende)

Dies ist ein Hinweis darauf, dass die Clusterbildung nach der 228. Stufe sinnvoll beendet werden sollte und die bis dahin gefundene Clustereinteilung das Endergebnis ist. Nach Kopp und Lois (2009, S. 49) kann die optimale Clusteranzahl nun abgelesen werden,

> *„indem man die Differenz zwischen der Anzahl der zu klassifizierenden Objekte [...] und dem Fusionsschritt bildet, nach dem ein deutlicher Anstieg der Distanzwerte ('Koeffizienten') zu beobachten ist [...]”.*

Betreffend der täglichen Fernsehdauer ist dementsprechend eine 4er-Clusterlösung, als Ergebnis der Differenz 232-228, optimal.

Die Spalten „Erstes Vorkommen des Clusters” und „Nächster Schritt” sind vor allem hinsichtlich des Überblicks von Bedeutung und sollen daher hier nicht näher beleuchtet werden.

Cluster-Zugehörigkeit

Fall	4 Cluster
1	1
2	2
3	1
4	2
5	2
6	2
7	2
8	1
9	3
10	2
11	2
12	4
13	2
14	1
15	1
16	1
17	4

Abbildung 6.9 Clusterzugehörigkeit

Die entscheidende Tabelle „Cluster-Zugehörigkeit” zeigt nun für jeden Probanden an, zu welchem der vier identifizierten Cluster er oder sie gehört. Die ersten acht Probanden werden beispielsweise in Cluster 1 oder 2 einsortiert, der neunte Proband in Cluster 3 und der zwölfte Proband in Cluster 4. Die Überschrift „4 Cluster” informiert darüber, dass der Datensatz in vier Gruppen eingeteilt wird.

Da alle Probanden einer der vier Gruppen zugeordnet werden können, lässt sich sowohl die genaue Verteilung als auch die jeweilige Zeitspanne des täglichen Fernsehkonsums, die die vier Gruppen repräsentieren, bestimmen. Die Probanden in den Clustern werden zu diesem Zweck fortan als **Wenigseher**, **Mittelseher**, **Vielseher** und **Exessivseher** interpretiert. Es folgt das erste fundamentale Auswertungsergebnis des Forschungsprojekts, dargestellt in der Tabelle „Die Gruppen der täglichen Fernsehdauer”.

Tabelle 6.2

Die Gruppen der täglichen Fernsehdauer

	Wenigseher	Mittelseher	Vielseher	Exzessivseher
Probanden	30,2 %	32,8 %	19,4 %	17,7 %
Zeitspanne	bis 1h 11min	1h 12min - 2h 15min	2h 16min - 3h 23min	ab 3h 24min

(i) **30,2 %** der Probanden sind **Wenigseher**, die maximal **1 Stunde 11 Minuten** pro Tag fernsehen.

(ii) **32,8 %** der Probanden sind **Mittelseher**, die zwischen **1 Stunde 12 Minuten** und 2 **Stunden 15 Minuten** pro Tag fernsehen.

(iii) **19,4 %** der Probanden sind **Vielseher**, die zwischen **2 Stunden 16 Minuten** und **3 Stunden 23 Minuten** pro Tag fernsehen.

(iv) **17,7 %** der Probanden sind Exzessivseher, die mindestens **3 Stunden 24 Minuten** pro Tag fernsehen.

Dieses Ergebnis lässt sich auch mit Hilfe vierer Boxplots darstellen:

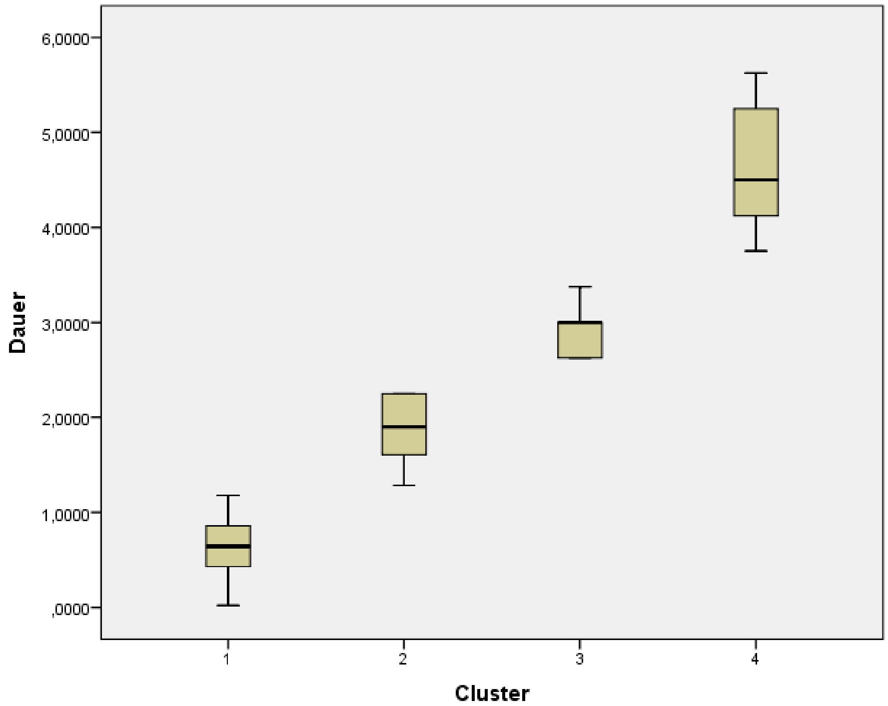


Abbildung 6.10 Boxplot Gruppen der Fernsehdauer

Das Diagramm illustriert die vier Cluster, wobei Cluster 1 in dieser Grafik über die *Wenigseher*, Cluster 2 über die *Mittelseher*, Cluster 3 über die *Vielseher* und Cluster 4 über die *Exzessivseher* Auskunft gibt. Da keine Ausreißer zu erkennen sind, repräsentieren jeweils die oberen und unteren Grenzen der Boxplots die oben aufgeführte Zeitspannen. Der Median beträgt bei den *Wenigsehern* **39 Minuten**, bei den *Mittelsehern* **1 Stunde 54 Minuten**, bei den *Vielseher* **3 Stunden** und bei den *Exzessivseher* unfassbare **4 Stunden 30 Minuten**.

Nachdem es gelungen ist, die tägliche Fernsehdauer der Gesamtpopulation der Stichprobe auf die vier Gruppen zu verteilen, lohnt zum Abschluss dieses Abschnitts ein kurzer Blick darauf, wie sich die Geschlechter und Schulformen auf die drei identifizierten Gruppen verteilen:

Tabelle 6.3

Gruppen täglicher Fernsehdauer und Geschlecht

	Wenigseher	Mittelseher	Vielseher	Exzessivseher
männlich	28,6 %	36,5 %	18,3 %	16,7 %
weiblich	32,1 %	28,3 %	20,8 %	18,9 %

Die Tabelle 6.3 zeigt dabei den Anteil der männlichen und weiblichen *Wenig-*, *Mittel-*, *Viel-* und *Exzessivseher*. Insgesamt verteilen sich sowohl die Schülerinnen als auch die Schüler recht ähnlich auf die vier Gruppen. Nichtsdestotrotz fällt die Häufung der weiblichen Probanden mit extremen Sehgewohnheiten auf. Dies meint, dass bei den Jungen deutlich mehr *Mittelseher* (36,5 %) zu identifizieren sind, als bei den Mädchen (nur 28,6 %), wohingegen die Mädchen in den extremen Kategorien (*wenig*, *exzessiv*) vor den Jungen liegen. Der Anteil der weiblichen *Viel-* und *Exzessivseher* ist überraschenderweise höher als der männliche Vergleichswert.

Tabelle 6.4

Gruppen täglicher Fernsehdauer und Schulform

	Wenigseher	Mittelseher	Vielseher	Exzessivseher
Gymnasium	51,1 %	32,2 %	10,0 %	6,7 %
Gesamtschule	18,3 %	38,5 %	23,1 %	20,2 %
Hauptschule	13,2 %	18,4 %	31,6 %	36,8 %

Wesentlich interessanter ist die Tabelle 6.4, welche sich mit den verschiedenen Schulformen hinsichtlich der Fernsehdauer beschäftigt. Dabei wird deutlich, dass der Großteil der Gymnasiasten (51,1 %) *Wenigseher* sind, während es kaum *exzessiv* fernsehende Jugendliche gibt (nur 6,7 %). An der Gesamtschule sind die meisten Jugendlichen *Mittelseher* (38,5 %), wobei es mehr *Viel-* und *Exzessivseher* als *Wenigseher* gibt. Noch erschreckender fällt das Hauptschulfazit aus. Lediglich 13,2 % der Schülerinnen und Schüler schauen *wenig* fern. Jeweils über 30 % dieser Jugendlichen nutzen das Fernsehen dagegen *viel* und *exzessiv*. Die meisten *Viel-* und *Exzessivseher* kommen demnach anteilsmäßig von der Hauptschule (insgesamt 68,4 %).

6.2.3 Zusammenfassung „Fernsehdauer und -häufigkeit"

In einem ersten Schritt ist es das Ziel, die tägliche Fernsehdauer der Probanden zu erheben. Dies geschieht auf Grundlage einer Fragebogenaktion, bei der die Schülerinnen und Schüler zunächst die offene Frage nach ihrer täglichen Fernsehdosis und im Anschluss die geschlossene Frage nach den bevorzugten Tagen des Fernsehkonsums beantworten. Anhand der erhobenen Daten ist es möglich, durch die Formel $\frac{Stundenzahl \cdot Tage}{7}$ die tägliche Fernsehdauer der Probanden zu berechnen. Es stellt sich heraus, dass die durchschnittliche Fernsehdauer der Schülerinnen und Schüler insgesamt **2 Stunden 13 Minuten** beträgt. Gymnasiasten schauen am wenigsten fern, während HauptschülerInnen am meisten fernsehen. Zwischen Jungen und Mädchen können keine signifikante Unterschiede festgestellt werden.

Mit Hilfe der Clusteranalyse (es ist eine begründete Entscheidung für den Ward-Algorithmus als Methode und die quadrierte Euklidische Distanz als Proximitätsmaß getroffen worden) gelingt es im Anschluss vier Gruppen unterschiedlicher täglicher Fernsehdauer zu ermitteln. Die größte Gruppe sind die *Mittelseher* (32,8 %), vor den *Wenigsehern* (30,2 %), den *Vielsehern* (19,4 %) und den *Exzessivseher* (17,7 %).

6.3 Die latenten Faktoren des Fernsehkonsums

Nachdem sich der erste Schritt der Studie mit der Fernsehdauer beschäftigt hat, folgt im Anschluss einer der zentralsten Aspekte dieser Arbeit: die Ermittlung der dem Fernsehverhalten der Jugendlichen zugrunde liegenden latenten Faktoren. Vereinfacht formuliert, soll untersucht werden, auf welche Arten die Probanden das Fernsehen nutzen und welcher Faktor besonders dominant ist, so dass die Schülerinnen und Schüler im Anschluss entsprechend ihres Antwortverhaltens auf die verschiedenen Nutzungsgruppen verteilt werden können. Somit wird es möglich sein, für jeden Probanden individuell den Anspruch und das Niveau des jeweiligen Fernsehprogramms anzugeben!

Um dieses Ziel zu erreichen, wird die folgende Vorgehensweise umgesetzt: Zunächst werden durch Schülerbefragungen diejenigen Fernsehsendungen ermittelt, die sich die Jugendlichen tatsächlich ansehen. Die durch diesen ersten Schritt erhaltene Liste wird anschließend der vollständigen Probandengruppe mit der Bitte, die aufgelisteten Fernsehsendungen in eine individuelle Prioritätsordnung zu überführen, vorgelegt. Da die subjektive Bewertung der Probanden nur ordinale Signifikanz besitzt, darf auch die daraus resultierende Rangordnung nur ordinale Signifikanz haben. Daher wird eine lückenfreie Anordnung verwendet, die aber verbundene, also mehrmalig gleich vergebene Ränge zulässt. Basierend auf dem dadurch gewonnenen Datensatz können die dem Fernsehverhalten der Jugendlichen zugrunde liegenden latenten Faktoren durch das im mathematischen Kapitel vorgestellte, an die übliche Faktorenanalyse angelehnte Verfahren ermittelt werden. Wie ausführlich dargestellt, ist das zu verwendende Verfahren, welches auf dem Einsatz der 0-1-Daten beruht, speziell auf Rangdaten zugeschnitten, wo die üblichen Verfahren der Faktorenanalyse versagen, so dass abschließend tatsächlich das Fernsehverhalten der befragten Probanden durch die gewonnenen latenten Faktoren beschrieben werden kann. Dieses Ergebnis soll dann im optimalen Fall durch das an die formale Begriffsanalyse angelehnte Verfahren bestätigt werden. Gelingt dies, wäre das Ergebnis methodenunabhängig, mit der zusätzlichen Konsequenz, dass der Informationsverlust der Faktorenanalyse tatsächlich, wie beabsichtigt, minimiert wird. Schließlich ist die formale Begriffsanalyse immun gegenüber einem solchen Informationsverlust, da keine Ähnlichkeitsmatrix dem Verfahren zugrunde liegt.

6.3.1 Datenerhebung

Egal wie ausgeklügelt die Vorgehensweise auch sein mag, zu lösende Probleme bzw. zu beantwortende Fragen treten nichtsdestotrotz auf. Solch ein Problem zeigt sich beispielsweise gleich zu Beginn bei der Auswahl derjenigen Fernsehendungen, die sich die Probanden tatsächlich ansehen.

Zunächst wird dazu eine vollständige Liste aller Sendungen, die im März 2011 aktuell oder regelmäßig im frei empfangbaren deutschen Fernsehen ausgestrahlt wurden, aufgestellt. Konsequenterweise werden dabei sowohl die öffentlich-rechtlichen Fernsehsender ARD, ZDF, 3SAT, ARTE, PHOENIX, KI.KA, WDR, BR, MDR, NDR und SWR als auch die privaten Kanäle RTL, RTL II, VOX, SUPERRTL, N-TV, PROSIEBEN, SAT.1, KABEL EINS, N24, 9LIVE, SIXX, TELE 5, VIVA, NICK, SPORT1, EUROSPORT und NRW TV berücksichtigt. Nun liefert aber die Recherche mit Hilfe der Zuschauerredaktionen sowie der Informationen über die jeweiligen Homepages im Internet (Sendungen A-Z) eine gigantische Anzahl an Sendungen, so dass die Liste 52 Seiten umfasst.

Natürlich ist eine solche Liste, man erinnere sich an das Motto, dass in der Fragebogenkürze die Ausfüllwürze liege (vgl. KIRCHHOFF et al. 2006, S. 27), für eine Fragebogenaktion viel zu umfangreich. Es steht außer Frage, dass es völlig sinnlos ist, 232 Probanden mit der Bitte, diejenigen Fernsehsendungen, die sie kennen, anzukreuzen, vor einen solchen Fragebogen zu setzen. Um aber dennoch auf wissenschaftlich vertretbare Art und Weise die Anzahl der Sendungen zu reduzieren, wird daher per Zufallsauswahl aus jeder teilnehmenden Klasse eine Handvoll Schülerinnen und Schülern ausgewählt, mit denen der Verfasser sich in Kleingruppengesprächen jeweils über sämtliche Fernsehsendungen auf der 52seitigen Liste austauschen kann. Mit Hilfe dieser *Experten*-Einschätzungen ist es möglich, die Liste um diejenigen Sendungen, die vollkommen unbekannt sind oder überhaupt nicht geschaut werden, zu kürzen. *Obwohl diese Arbeit für die Jugendlichen nicht einfach und durchaus fordernd gewesen ist, haben sämtliche Probanden hoch motiviert mitgeholfen, so dass von einem absolut brauchbaren Ergebnis ausgegangen werden kann.* Zwar bleiben auch nach diesem Arbeitsschritt noch zahlreiche Sendungen übrig, doch ist nun ein Umfang erreicht, der einen Fragebogen legitimiert, den man guten Gewissens allen 232 Schülerinnen und Schülern vorlegen kann, um Schritt 1, die Auswahl der tatsächlich geschauten Sendungen, abzuschließen.

Der entsprechende Fragebogen, dessen Deckblatt und eine weitere Beispielseite unten abgebildet sind, beginnt mit einer kurzen Erklärung des Forschungsprojekts für die Probanden. Nach Bühner ist ein solcher Einstieg zwingend notwendig, um die teilnehmenden Schülerinnen und Schüler für die Thematik zu sensibilisieren (vgl. Bühner 2006, S. 69). Die Einführung endet mit der konkreten Aufgabenstellung: *„Deine Aufgabe ist es heute, die Sendungen anzukreuzen, die du tatsächlich oft (bzw. öfters) schaust."*

UNIVERSITÄT DUISBURG ESSEN

FAKULTÄT FÜR MATHEMATIK
MICHAEL FEY
PROF. DR. GERHARD HERDEN

Fragebogen zum Thema Fernsehsendungen

http://witze.net/mann,steinzeit,fernsehen.gif

ARD

Sendung	
Anne Will	
ARD Fernsehlotterie	
ARD-Buffet	
ARD-exclusiv	
ARD-Mittagsmagazin	
ARD-Wetterschau	
BAMBI (Musik-Preisverleihung)	
Blaubär und Blöd	
Börse im Ersten	
Brisant	
Das Duell im Ersten	
Das fantastische Quiz des Menschen	
Das Hochzeitsschiff	
Das Quiz der Deutschen	
Das unglaubliche Quiz der Tiere	
Das Wetter im Ersten	
Das Wort zum Sonntag	
Der Dicke	
Der Wolf	
Deutschland, deine Künstler	
Die Anwälte	
Die große Show der Naturwunder	
Die große Wissensshow mit Ranga Yog	
Die großen Kriminalfälle	
Die Kinder vom Alstertal	
Die Pfefferkörner	

Sendung	
Eurovision Song Contest	
Fakt	
Felix und die wilden Tiere	
Feste der Volksmusik	
Frag doch mal die Maus	
Garfield	
Giraffe, Erdmännchen & Co.	
Gott und die Welt	
Großtstadtrevier	
Hakan Nesser	
Harald Schmidt	
Hart aber Fair	
Im Angesicht des Verbrechens	
Immer wieder sonntags	
In aller Freundschaft	
Jim Knopf	
Karen in Action	
Kopfball	
Lindenstraße	
Lotto-Ziehung	
Mama ist unmöglich	
Märchenfilme	
Marienhof	
Meine peinlichen Eltern	
Menschen bei Maischberger	
Mitgemacht und mitgelacht	

Abbildung 6.11 Fragebogen Deckblatt und Seite 1

Nach Durchführung dieses Schrittes ist es möglich, die meistgeschauten Sendungen statistisch festzuhalten, so dass das Ergebnis der Auswertung offenkundig zeigt, welche Sendungen von den Jugendlichen vorwiegend geschaut werden. Um nun die Liste mit den meistgeschauten Sendungen fertigzustellen, muss die Entscheidung getroffen werden, wie viele Schülerinnen und Schüler eine Sendung oft bzw. öfters schauen müssen, damit das Format in die finale Liste aufgenommen werden kann. Dabei konkurrieren zwei Überlegungen: Einerseits sollen möglichst viele Sendungen, um hinsichtlich der späteren Faktorenanalyse mit hochwertigen Ergebnissen rechnen zu können, ausgewählt werden. Andererseits macht es keinen Sinn, Sendungen aufzunehmen, die nur von sehr wenigen Probanden geschaut werden. Daher entscheidet der Verfasser nach eigenem Ermessen, dass eine Fernsehsendung genau dann in die finale Liste der meistgeschauten Sendungen aufgenommen wird, wenn mindestens 25 % der Probanden diese oft bzw. öfters schauen. Nach diesem Vorgehen verfahrend, können letztlich genau 80 Fernsehsendungen in die Liste für den zweiten Schritt und somit den zweiten Fragebogen aufgenommen werden.

Wenngleich die folgende Statistik nach Ablauf des ersten Schrittes noch nicht besonders aussagekräftig ist und maximal eine tendenzielle Auskunft über die Sehgewohnheiten der Schülerinnen und Schüler gibt, können bereits zu diesem Zeitpunkt einige interessante Entdeckungen in einem kurzen Exkurs präsentiert werden.

Tabelle 6.5

Die Top 10 der tatsächlich geschauten Sendungen

Sendung	Sender	Stimmen
Das Supertalent	RTL	179
Deutschland sucht den Superstar	RTL	174
Galileo	Pro7	168
Die Simpsons	Pro7	167
Schlag den Raab	Pro7	159
Familien im Brennpunkt	RTL	146
Verdachtsfälle	RTL	146
Betrugsfälle	RTL	139
Schulermittler	RTL	137
Family Guy	VIVA	128

Tabelle 6.5 zeigt die zehn Fernsehsendungen, die von den Probanden die meisten Stimmen bekommen haben. Es handelt sich somit um diejenigen Sendungen, die bei Jugendlichen am bekanntesten sind. Es fällt auf, dass die öffentlich-rechtlichen Fernsehsender in dieser Liste vollständig fehlen. Das alltägliche Fernsehprogramm der Probanden wird offensichtlich stark von den privaten TV-Sendern dominiert. Hervorzuheben ist, dass mit *Galileo* zumindest eine Sendung auf den vorderen Rängen vertreten ist, die immerhin den Anspruch verfolgt, Wissen, Bildung und Informationen zu liefern. Aufgrund seiner Bauart lässt sich die Sendung zwar eher als Infotainmentmagazin denn als Wissenschaftsmagazin einsortieren, dennoch sticht die Sendung sicherlich aus der Liste der bekanntesten Formate heraus. Während auch die beiden Zeichentrickserien *Die Simpsons* und *Family Guy* zu den intelligenteren, mit zahlreichen Verweisen und Metaphern gespickten Formaten im Fernsehen gezählt werden dürfen und die Mischung aus Quiz und Sport *Schlag den Raab* ebenfalls zu einer weitestgehend ambitionierten Fernsehsendung werden lässt, muss allerdings genauso festgehalten werden, dass mehr als die Hälfte der Top 10-Sendungen umgangssprachlich als „Trash"-TV oder wahlweise als „Hartz IV"-TV bezeichnet werden. Gemeint ist damit das fehlende inhaltliche und schauspielerische Niveau von Reality-Dokus wie *Familien im Brennpunkt, Verdachtsfälle, Betrugsfälle* und *die Schulermittler* oder den RTL-Castingshows *Das Supertalent* und *Deutschland sucht den Superstar.*

Nach diesem kurzen Abstecher in den Bereich der ersten kleinen Forschungsergebnisse basierend auf der Ermittlung derjenigen Fernsehsendungen, die sich die Jugendlichen tatsächlich ansehen, soll nun wieder die Bestimmung der latenten Faktoren des Fernsehkonsums im Zentrum der Überlegungen stehen. Dazu gilt es nun den zweiten Fragebogen auf Basis der in Schritt 1 gewonnenen 80 Sendungen aufzustellen, damit die Probanden diese 80 Sendungen in ihre jeweilige individuelle Prioritätsordnung überführen können. Vereinfacht formuliert ist es im zweiten Schritt die Aufgabe der Jugendlichen, die 80 Sendungen hinsichtlich der jeweiligen Beliebtheit zu ordnen. Diese Aufgabe bringt allerdings ein weiteres zu lösendes Problem mit sich. Es stellt sich nämlich die berechtigte Frage, wie die Probanden in die Lage versetzt werden sollen, 80 Sendungen anordnen zu können.

Für dieses Problem ist eine spielerische Lösung gefunden worden, indem die Jugendlichen in die Rolle von Fernsehtestern wechseln dürfen, die die Aufgabe bekommen, die Fernsehsendungen mit Zensuren zu bewerten. Aus eben diesen Zensuren ist es dann für den Verfasser möglich, die jeweilige individuelle Prioritätsordnung der Probanden zu rekonstruieren. Dafür ist es aber von entscheidender Bedeutung, dass die Schülerinnen und Schüler mindestens 80 Zensuren verwenden dürfen, denn nur somit ist gewährleistet, dass ein Proband zumindest theoretisch jeder Sendung eine individuelle Note zuweisen kann.

Die Rollen- und damit die Arbeitsanweisung für die Probanden lautet:

> *„Du schaust regelmäßig Fernsehen? Du hast eine absolute Lieblingssendung? Oder geht dir das TV-Programm gehörig auf die Nerven? Gut! Denn heute wirst Du zum Fernseh-Tester! Als Fernseh-Tester ist es deine Aufgabe, TV-Sendungen zu benoten! Dazu musst du natürlich nicht jede Sendung regelmäßig oder besonders oft gucken. Allerdings solltest du wissen, worum es ungefähr geht, damit du sie bewerten kannst. Wenn du eine TV-Sendung gar nicht kennst, kannst du auch dies ankreuzen. Dir stehen dabei aber nicht nur die 6 Schulnoten zur Verfügung, sondern noch viele mehr. Du darfst nämlich für jede Sendung eine Note zwischen 1 (Lieblingssendung) und einer beliebig großen Zahl (mag ich gar nicht) vergeben!"*

Die Gestaltung des Deckblatts und des eigentlichen Fragebogens ist unten abgebildet:

TV-Sendung	Note	Kenne ich nicht
Alarm für Cobra 11 (RTL)		
Betrugsfälle (RTL)		
Big Brother (RTL II)		
Big Time Rush (Nickelodeon)		
Boxen bei RTL (RTL)		
Bundesliga Aktuell (Sport 1)		
Comedystreet (Pro7)		
Cosmo und Wanda (Nickelodeon)		

Abbildung 6.12 Finaler Fragebogen

Nun mag es sicherlich Einwände geben, dass der gleiche Fragebogen gleichen Probanden an verschiedenen Tagen vorgelegt zu völlig unterschiedlichen Bewertungen führen kann. Letztlich ist diese Problematik bei einer derart komplexen Aufgabe für die Probanden sicherlich nicht vollkommen auszuräumen, doch geht der Verfasser davon aus, dass die angeführte Problematik vor allem darin begründet ist, dass die Reflexion der Alternativen durch die Probanden nicht ausreichend, weil in der Regel zu kurz ist. Daher ist es nach Meinung des Verfassers dringend nötig, der Datenerhebungsphase eine intensive Reflexionsphase der Alternativen durch die Probanden vorausgehen zu lassen, in der jeder Proband einen Ansprechpartner hat, der ihm hilft, vollständige Klarheit über die zu bewertenden Alternativen zu gewinnen. Dadurch kann die Problematik sicherlich abgeschwächt werden, so dass wenigstens ordinale Signifikanz der Bewertungen durch die Probanden vorausgesetzt und auf Grundlage der Überlegungen des mathematischen Kapitels die verifizierte Form der Faktorenanalyse zur Auswertung benutzt werden kann. Aus diesem Grund ist der Erhebungsphase eine aus-führliche Reflexionsphase vorangestellt, die die Schülerinnen und Schüler auf ihre Aufgabe und die zu bewertenden Sendungen vorbereitet.

6.3.2 Umgang mit dem Datensatz

Im Anschluss an die von den Schülerinnen und Schülern vorgenommene Notenvergabe können die Zensuren vom Verfasser in die benötigten Ränge übertragen werden, damit für jeden Probanden die individuelle Prioritätsordnung seines Fernsehprogramms entsteht. In Anlehnung an die in Sportwettkämpfen zu erzielenden Rangordnungen wird dem Rangplatz 1 fortan die oberste Probandenpräferenz und dem Rangplatz 80 die niedrigste Probandenbeliebtheit zugeschrieben. Durch dieses Vorgehen kann folgender Datensatz generiert werden, der in Abbildung 6.13 auszugsweise abgebildet ist:

	P001	P002	P003	P004	P005	P006	P007	P008	P009	…	P231	P232
Sendung 1	36	35	1	9	9	21	30	22	75	…		10
Sendung 2	37	37	42	21		19	55	70	11	…		40
Sendung 3	67		41				78	76	80	…		80
Sendung 4	44	32	43		11		10		63	…		50
Sendung 5	38		13	3	18	9	71	21	35	…	79	50
Sendung 6	45		16	6	4			77	34	…	40	50
Sendung 7	14	20	15	69			60	63	42	…		
Sendung 8	28	39	40	51		53	20	49	33	…	30	40
Sendung 9	39	9			45	56	21	19	79	…		3
Sendung 10	6		12	16	7	58	11	24	10	…	80	20
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Sendung 79	22	24	33	5	21	60	41	15	57	…	80	40
Sendung 80	23	52	36	10	73	69	42	17	19	…		15

Abbildung 6.13 Lückenhafter Datensatz zur Bestimmung der latenten Faktoren des Fernsehkonsums

In der obersten Zeile der Tabelle sind jeweils die Probanden von P001 bis P232 angegeben und in der ersten Spalte die 80, in Schritt 2 herausgefilterten Fernsehsendungen. Zu jeder TV-Sendung ist der von jedem Probanden individuell vergebene Rangplatz einzusehen.

Beim Betrachten des Datensatzes fällt auf, dass zahlreiche Bewertungen fehlen. Da auf Basis der Auswahlkriterien aus Schritt 1 bekannt ist, dass einige der 80 Sendungen nur von 25 % der Schülerinnen und Schüler häufig geschaut werden, musste der Verfasser bereits im Vorfeld darauf gefasst sein, dass ein bestimmter Anteil der Sendungen von den Probanden nicht bewertet werden kann. Die Anzahl der fehlenden Werte beträgt nun aber 29 %, was einem sehr hohen Wert entspricht. Selbst wenn es gelingt, durch Einsatz einer adäquaten und funktionalen Methode die *missing values* so zu rekonstruieren, dass der Datensatz vollständig aufgefüllt werden kann, besteht aufgrund der Vielzahl an fehlenden Werten das Risiko die ursprünglich so penibel erhobenen Daten zu verfälschen. Daher werden diejenigen Fernsehsendungen eliminiert, die von weniger als 70 % der Probanden bewertet wurden. Anhand dieses Vorgehens, bei dem sich der Datensatz auf 53 Sendungen reduziert, gelingt es selten geschaute Sendungen auszuschließen und einen zu über 80 % vollständigen Datensatz zu generieren. Die übrig gebliebenen fehlenden Werte sollen nun entsprechend des aus dem mathematischen Kapitel bekannten Verfahrens geschätzt werden.

6.3.3 Die Ermittlung der latenten Faktoren des Fernsehkonsums

Nach der Schätzung sämtlicher *missing values* liegt nun ein vollständiger Datensatz für 53 Sendungen vor:

	P001	P002	P003	P004	P005	P006	P007	P008	P009	…	P231	P232
Alarm für Cobra 11	27	25	1	7	9	17	23	18	49	…	11	3
Betrugsfälle	28	27	36	16	24	15	32	46	11	…	11	8
Big Brother	50	39	35	40	38	37	41	50	52	…	11	14
Boxen bei RTL	29	14	11	3	15	7	38	17	26	…	10	9
Bundesliga Aktuell	35	39	13	6	5	13	9	51	25	…	7	9
Comedystreet	13	18	12	38	3	4	33	40	30	…	11	4
Cosmo und Wanda	23	28	33	32	35	28	17	35	24	…	6	8
CSI: Miami	30	10	30	18	23	30	18	15	51	…	11	2
Das Supertalent	5	14	10	13	7	33	12	19	10	…	12	5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	…	…
X-Factor: Das Unfassbare	42	17	50	27	19	6	27	13	16	…	11	9
Zack & Cody an Board	18	19	27	5	17	34	30	14	38	…	12	8

Abbildung 6.14 Vollständiger Datensatz zur Bestimmung der latenten Faktoren des Fernsehkonsums aus 53 Sendungen

Bevor nun mit der Faktorenanalyse (hier der Hauptkomponentenmethode mit anschließender Varimax-Rotation) begonnen wird, werden die Ränge entsprechend der Diskussion im mathematischen Kapitel dichotomisiert bzw. in 0-1-Daten übertragen.

Die wichtigsten Grundlagen der Faktorenanalyse sind bereits im eben zitierten vierten Kapitel dargestellt worden, daher soll an dieser Stelle parallel zur Extraktion der latenten Faktoren des Fernsehkonsums einmalig auf die gewählte Extraktionsmethode, die Eignung der Daten für eine Faktorenanalyse, verschiedene Rotationsmethoden und ihre Funktionen, zwei Kriterien zur Bestimmung der Faktorenanzahl sowie die Interpretation und Benennung der Faktoren eingegangen werden. All diese Aspekte werden bei den weiteren analog verwendeten Faktorenanalysen dieser Arbeit als bekannt vorausgesetzt.

Extraktionsmethode

In dieser Arbeit soll bei sämtlichen Faktorenanalysen die Hauptkomponentenmethode (Principal Component Analysis; PCA) Anwendung finden. Ziel der PCA ist es, die maximale Aufklärung der Gesamtvarianz aller manifesten Variablen in der Rohdatenmatrix durch möglichst wenige Komponenten zu erreichen (vgl. Schendera 2010, S. 191). Vereinfacht ausgedrückt soll die PCA zu einer Datenreduktion und einer Datenerklärung beitragen. Aus diesem Grund werden bei der PCA lineare Kombinationen der Variablen gebildet, so dass als erste Hauptkomponente bzw. als erster Faktor derjenige ausgewiesen wird, der den größten Teil der Gesamtstreuung aller Variablen im statistischen Sinne ergibt, so dass zueinander orthogonale Komponenten die vollständige Varianz der Rohdatenmatrix erklären. Der erste Faktor wird über eine Linearkombination von Variablen bestimmt, die ein Maximum der Varianz erklären, der zweite Faktor erklärt dementsprechend ein Maximum der Restvarianz usw (vgl. ebd., S. 194). Da die PCA, wie im mathematischen Kapitel beschrieben, dabei von einer Korrelationsmatrix ausgeht, wird bei Rangdaten eine Transformation der ordinal skalierten Daten in dichotomisierte 0-1-Daten notwendig um einen nicht zu verantwortenden Informationsverlust auszuschließen.

Eignung der Daten für eine Faktorenanalyse

Mit der Wahl der Extraktionsmethode ist die Arbeit allerdings noch längst nicht getan. Auch bei der Durchführung und Auswertung gilt es weitere Problembereiche zu beachten, um zu einem mathematisch vertretbaren Ergebnis zu gelangen. Zunächst müssen die Daten sich überhaupt zur Durchführung einer PCA eignen, da die Faktorenanalyse voraussetzt, dass die Variablen hoch genug miteinander interkorrelieren. Daher wird ein Verfahren benötigt, welches die Daten einer generellen Prüfung unterzieht (vgl. Klopp o.J.). Führt man die Faktorenanalyse nun in SPSS durch, wird aus diesem Grund zunächst der Kaiser-Meyer-Olkin-Test (KMO) durchgeführt, welcher die Größe der partiellen Korrelationen zwischen den Variablen prüft (vgl. Schendera 2010, S. 245). Der KMO-Wert errechnet sich anhand der Formel:

$$KMO = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij}^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij}^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij.z}^2} (i \neq j)$$ (Bühner 2006 zitiert nach Klopp o.J.)

KMO- und Bartlett-Test

Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.		,984
Bartlett-Test auf Sphärizität	Ungefähres Chi-Quadrat	589908,845
	df	1378
	Signifikanz nach Bartlett	,000

Abbildung 6.15 KMO- und Bartlett-Test (latente Faktoren des Fernsehkonsums)

Für den, den Fernsehkonsum betreffenden Datensatz ergibt sich nach dieser Berechnung ein KMO-Wert von 0,984, so dass sich die erhobenen Daten in besonderem Maße für eine Faktorenanalyse eignen, da nach Kaiser und Rice (1974) folgende Abstufung gilt:

KMO-Koeffizient	Eignung der Daten
>0,90	sehr gut
0,80-0,90	gut
0,70-0,79	mittel
0,60-0,69	mäßig
0,50-0,59	schlecht
<0,50	inkompatibel mit der Berechnung.

Diese Einteilung lässt sich damit begründen, dass ein großer KMO-Koeffizient bedeutet, dass die Summe der Partialkorrelationskoeffizienten klein ist und damit die Interkorrelationsmatrix viel gemeinsame Varianz enthält (vgl. Klopp o.J.).

Anzahl der zu extrahierenden Faktoren

Nachdem die Eignung des Datensatzes sichergestellt ist, stellt sich die elementare Frage nach der Anzahl der zu extrahierenden Faktoren. Dabei gilt es grundsätzlich zu beachten, dass ein hinreichend großer Teil der Streuung durch die Faktoren erklärt werden muss, da diese schließlich das Abbild der Ausgangsdaten sein sollen. Anderseits sollen die Faktoren natürlich auch zu einer ausreichend großen Reduzierung der Komplexität beitragen. Eine Faktorenanalyse verfolgt grundsätzlich immer das Ziel, die „richtige" Anzahl an Faktoren zu extrahieren. In der Literatur werden einige Verfahren, sogenannte Abbruchkriterien, genannt, mit deren Hilfe die Anzahl der Faktoren bestimmt werden kann. Besonders häufig wird entweder der Scree-Test von Cattell (1966) oder Kaisers Eigenwertkriterium (1970) angewendet.

(i) Der Scree-Test ist eine Visualisierung der anfänglich nacheinander ermittelten Eigenwerte und der Anzahl der zu extrahierenden Faktoren in einem zweidimensionalen Koordinatensystem. Die Eigenwerte und Faktoren beginnen dabei mit einer abfallenden Kurve, die meist einen auffälligen Knick aufweist. Nach diesem Knick verläuft die Kurve nahezu parallel zur x-Achse. Es wird darauf verwiesen, dass die Knickstelle Auskunft über die zu extrahierenden Faktoren gibt. In der Literatur ist allerdings umstritten, ob der Knickpunkt selbst, der Faktor davor oder der Faktor nach dem Knick die zu wählende Marke ist. Dementsprechend unterliegt der Scree-Test einer gewissen Subjektivität (vgl. SCHENDERA 2010, S. 211).

(ii) Das bekannteste Verfahren ist das Eigenwertkriterium nach KAISER. Hierbei gilt es, diejenigen Faktoren zu extrahieren, deren Eigenwert größer als 1 ist, da alle Faktoren, die einen Eigenwert kleiner als 1 aufweisen, weniger erklären, als jede einzelne Variable selbst (vgl. KLOPP o.J.). Im weiteren Verlauf der Arbeit soll dieses Kriterium aufgrund seiner höheren Objektivität Anwendung finden, wenngleich auch dieses Kriterium in der Literatur durchaus umstritten ist.

Für den vorliegenden Datensatz liefert die Hauptkomponentenmethode dabei die Tabelle „Erklärte Gesamtvarianz". Diese gibt für die erzielte Lösung Eigenwerte, erklärte Varianz und erklärte kumulierte Varianz an.

Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	29,131	54,964	54,964	29,131	54,964	54,964	10,180	19,208	19,208
2	2,073	3,911	58,875	2,073	3,911	58,875	8,307	15,673	34,881
3	1,414	2,669	61,543	1,414	2,669	61,543	6,573	12,401	47,282
4	1,157	2,182	63,726	1,157	2,182	63,726	6,108	11,524	58,806
5	1,044	1,970	65,696	1,044	1,970	65,696	3,652	6,890	65,696
6	,932	1,759	67,455						
7	,844	1,592	69,047						

Abbildung 6.16 Erklärte Gesamtvarianz (latente Faktoren des Fernsehkonsums)

Die linke Hälfte der Tabelle bezieht sich auf die Anfangssituation der Faktorenextraktion, bei der so viele Faktoren wie Variablen vorhanden sind. Der Eigenwert eines Faktors gibt an, welcher Betrag der Gesamtstreuung durch diesen Faktor erklärt wird. Die rechte Hälfte der Tabelle stellt die Situation nach der Extraktion bzw. nach der Rotation dar. Es zeigt sich, dass der erste Faktor äußerst dominant ist und bereits 54,964 % der Gesamtvarianz erklärt. Auch die Faktoren zwei bis fünf besitzen einen Eigenwert größer als 1 und erklären weitere 10,732 % der Gesamtvarianz.

Somit sind nach dem Eigenwertkriterium von KAISER fünf Faktoren, die gemeinsam 65,696 % der Gesamtdaten erklären, zu extrahieren. Der resultierende Informationsverlust ist der Reduktion von 53 auf fünf Faktoren geschuldet, hält sich aber in einem mehr als vertretbaren Rahmen.

Interpretation der gefundenen Faktoren

Grundsätzlich ist ein latenter Faktor gemäß der Faktorentheorie keine Variable und kein Parameter der Stichprobe, sondern ein Populationsparameter bzw. ein invariantes Merkmal der Grundgesamtheit (vgl. SCHENDERA 2010, S. 212), welches es zu beschreiben gilt. Allerdings sind die im zweiten Schritt gefundenen Faktoren häufig zunächst schwierig zu interpretieren. Um die Interpretation zu erleichtern, macht man sich daher den Umstand zunutze, dass die Faktoren Gebilde sind, die sich verzerrungsfrei so transformieren lassen, dass sie in unterschiedlichen Koordinatensystemen dargestellt werden können. Durch eine geeignete Transformation gelingt es häufig, die Verbindungen zu den Beobachtungsvariablen deutlicher aufzuzeigen und damit die Interpretation zu vereinfachen. Eben dieser Schritt nennt sich Rotation (vgl. BROSIUS 1998, S. 643). Grundsätzlich können bei den Rotationsmethoden schiefwinklige und orthogonale Rotationsmethoden unterschieden werden. Da schiefwinklige Rotationsmethoden nur dann Verwendung finden können, wenn sichergestellt ist, dass die Faktoren miteinander korrelieren, sollen in diesem Forschungsprojekt orthogonale Rotationen, die auf Basis der gegenteiligen Annahme fußen, zur Erleichterung der Interpretation benutzt werden (vgl. SCHENDERA 2010, S. 205). Die Varimax-Rotation eignet sich dazu besonders gut. Bei dieser Methode wird die Anzahl von Variablen mit hoher Faktorenladung minimiert (vgl. KLOPP o.J.). Vereinfacht ausgedrückt bedeutet dies, dass pro Faktor einige Variablen hoch, alle übrigen aber möglichst gering laden sollen (vgl. SCHENDERA 2010, S. 206). Basierend auf dieser Methode, stellt SPSS dann die „rotierte Komponentenmatrix" (Abbildung 6.17) dar, durch welche die jeweiligen Variablen den fünf Faktoren zugeordnet werden.

Orientiert man sich an der Faustregel von GORSUCH (1983), sollen diejenigen Variablen zur Interpretation eines Faktors herangezogen werden, die betragsmäßig größer als 0,30 sind. Dies trifft in diesem Datensatz auf alle Variablen zu. Auch der differenzierte Interpretationshinweis von GUADAGNOLI und VELICER (1988) ist erfüllt. Dieser Ansatz meint, dass wenn auf einen Faktor mindestens vier Variablen Ladungen über 0,60 aufweisen, die Faktorenstruktur ungeachtet der Stichprobengröße generalisierend interpretiert werden kann (vgl. KLOPP o.J.).

Rotierte Komponentenmatrix[a]

	Komponente				
	1	2	3	4	5
Die Zauberer von Waverly Place	,715				
iCarly	,689				
Hotel Zack & Cody	,688				
DSDS- Das Magazin	,664				
Let's Dance	,650				
GZSZ	,648				
Big Brother	,643				
Zack & Cody an Board	,641				
Cosmo und Wanda	,610				
Mr. Bean	,580	,403			
Frauentausch	,575				
Mein Mann kann	,547	,482			
GNTM	,530				
Die Comedy Falle	,507				
Wetten, dass...?	,503	,491			
Taff	,496				
Ich bin ein Star - holt mich hier raus	,475	,406			
K11	,427				
Fußball live		,708			
Schlag den Raab		,664			
Boxen bei RTL		,636			

Abbildung 6.17 Rotierte Komponentenmatrix (latente Faktoren des Fernsehkonsums)

Die latenten Faktoren des Fernsehkonsums

Nach Durchführung sämtlicher relevanter Schritte und Überlegungen konnten die von den Schülerinnen und Schülern bewerteten Fernsehsendungen auf fünf Faktoren verteilt werden. Die entsprechende Aufteilung und die daraus resultierende Bezeichnung der Faktoren sieht folgendermaßen aus:

(i) *gemischte Unterhaltung:* Im ersten Faktor befinden sich verschiedene Unterhaltungssendungen, von denen einige einen relativ geringen Anspruch (*Big Brother*, *Ich bin ein Star*, *Frauentausch*) besitzen, während andere altersadäquate Unterhaltung bieten (*Hotel Zack & Cody*, *Mr. Bean*, *iCarly*). Zusammenfassend können diese Sendungen als durchschnittliches, **gemischtes Unterhaltungsfernsehen** interpretiert werden. Probanden, welche diesen Faktor bevorzugen, schauen reine Unterhaltung mit wechselndem inhaltlichen Anspruch.

(ii) *Wettkampf und Bildung:* Der zweite Faktor besteht (fast) vollständig aus Fernsehsendungen, die den Zuschauer entweder dazu befähigen, Wissen zu erwerben (bzw. sich bilden zu lassen) oder einem Wettkampf beizuwohnen. Wissen sei dazu in dieser Arbeit in Anlehnung an das Allgemeinkonzept von Hans Werner Heymann (1996) definiert. Heymann (1996, S. 79) sieht die Aufgabe der Schule darin, *„die Jugend mit Wissen über die Welt auszustatten"*. Dementsprechend sollen Schülerinnen und Schüler einen orientierenden Überblick über die Welt erhalten, damit sie die *„Erscheinungen um sich herum einzuordnen wissen"* (ebd.,

S. 79) und ihren eigenen Erfahrungshorizont erweitern können, um über die Welt Bescheid zu wissen. Somit dient Wissen der Weltorientierung, denn durch *„ein gewisses 'Luxurieren' des Wissens"* (ebd., S. 80) kann ein differenziertes Weltbild erschlossen und die eigene Stellung in der Welt eingeordnet werden. Eine Fernsehsendung die Wissen fördert, dient folgerichtig der Weltorientierung, dem Weltverständnis und dem Sehen-können der Welt. Sendungen, die diesen Wissensaspekt beinhalten sind *Galileo*, *Galileo Mystery* und *Nachrichtensendungen.* Die Wettkämpfe untergliedern sich in sportliche (*Fußball*, *Boxen*), musikalische (*Eurovision Song Contest*) oder spielerische (*Wer wird Millionär*, *Schlag den Raab*, *Schlag den Star*) Fernsehsendungen. Ausreißer innerhalb dieses Faktors sind zwar vorhanden (*Comedystreet*, *CSI: Miami*, *Alarm für Cobra 11*), aber bei der Menge an Variablen auch nicht völlig zu vermeiden, so dass der Faktor schlussendlich als **Wettkampf und Bildung** bezeichnet werden soll.

(iii) *niveauarme Unterhaltung:* Der dritte Faktor beinhaltet praktisch das komplette RTL-Programm. Einerseits sind hier die Doku-Soaps Verdachtsfälle, *Familien im Brennpunkt*, *Betrugsfälle* und *Die Schulermittler* beheimatet, andererseits lassen sich ebenso die viel kritisierten Castingshows *Das Supertalent* und *Deutschland sucht den Superstar* in diesem Faktor finden. Auch die übrigen Sendungen sind nicht gerade für anspruchsvolle Unterhaltung bekannt: Dieser Faktor kann guten Gewissens als **niveauarme Unterhaltung** bezeichnet werden.

(iv) *humorvolle Unterhaltung:* Im vierten Faktor befinden sich verschiedene Comedy-Sendungen. Dabei handelt es sich zum einen um lustige Zeichentrickserien (*Family Guy*, *Die Simpsons*, *Futurama*, *Spongebob*) und zum anderen um amerikanische Sitcoms (*Two and a half Men*, *How I met your Mother*, *Scrubs*). Daher soll dieser Faktor mit der Bezeichnung **humorvolle Unterhaltung** angegeben werden.

(v) *musikalische Unterhaltung:* Besonders leicht fällt die Interpretation des fünften Faktors, enthält dieser lediglich Musiksendungen. Folgerichtig wird er daher als **musikalische Unterhaltung** bezeichnet.

Die Verteilung der Fernsehsendungen auf die fünf extrahierten Faktoren, welche in der nachfolgenden Tabelle 6.6 detailliert für sämtliche Sendungen festgehalten wird, führt dementsprechend zu einem überraschend zufriedenstellenden Ergebnis:

Tabelle 6.6

Latente Faktoren des Fernsehkonsums (53 Sendungen)

gemischte Unterhaltung	Wettkampf und Bildung	niveauarme Unterhaltung	humorvolle Unterhaltung	musikalische Unterhaltung
Die Zauberer von Waverly Places	Fußball live	Verdachtsfälle	Family Guy	VIVA Top 20
iCarly	Boxen bei RTL	Familien im Brennpunkt	Two and a half Men	VIVA Top 100
Hotel Zack & Cody	Bundesliga Aktuell	Betrugsfälle	How I met your Mother	VIVA Charts
DSDS: Das Magazin	Schlag den Raab	Die Schulermittler	Die Simpsons	
Let's Dance	Schlag den Star	X-Diaries	Scrubs	
Gute Zeiten, schlechte Zeiten	Wer wird Millionär?	Vermisst	Futurama	
Big Brother	Eurovision Song Contest	Deutschland sucht den Superstar	Switch Reloaded	
Zack & Cody an Board	Elton vs. Simon	X-Factor: Das Unfassbare	Spongebob	
Cosmo und Wanda	Galileo Mystery	Das Supertalent		
Mr. Bean	Galileo			
Frauentausch	Nachrichten			
Mein Mann kann				
Germany's next Topmodel	**Ausreißer:**			
Die Comedy-Falle	Comedystreet			
Wetten, dass...?	CSI: Miami			
Taff	Alarm für Cobra 11			
Ich bin ein Star				
K11				
Willkommen bei Mario Barth				

Durch die Unterteilung der 53 Fernsehsendungen in fünf gut zu interpretierende Faktoren ist ein Hauptziel des Forschungsprojekts erreicht. Es ist gezeigt worden, dass Jugendliche fünf verschiedene Arten von Sendungen schauen, die die komplette Bandbreite der Unterhaltung abdecken.

Um nun zu überprüfen, inwieweit dieses Ergebnis methodenunabhängig ist, soll die ebenfalls im mathematischen Kapitel erläuterte modifizierte Version der formalen Begriffsanalyse, die einen durch die Faktorenanalyse möglichen Informationsverlust a priori ausschließt, auf den Datensatz angewendet werden. Diese Variante bezieht sich nur noch auf diejenigen Begriffe, die von einzelnen Sendungen erzeugt werden.

Das Ergebnis der formalen Begriffsanalyse wird in dem nachfolgendem Liniendiagramm dargestellt:

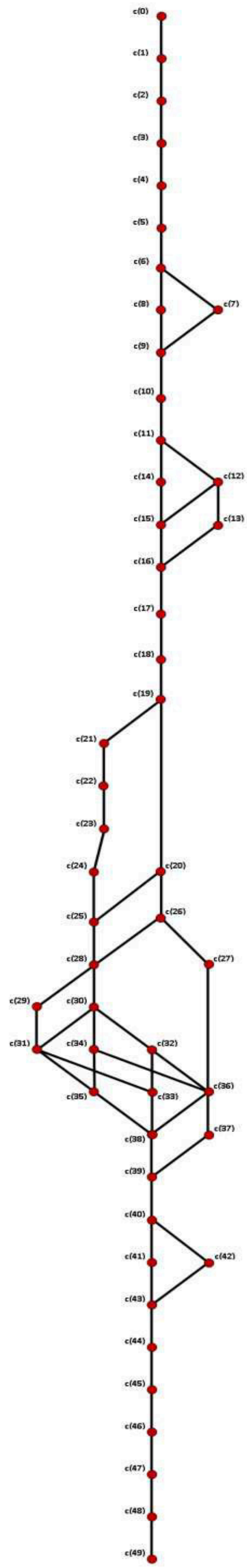


Abbildung 6.18 Liniendiagramm der formalen Begriffsanalyse (latente Faktoren des Fernsehkonsums)

Visuell beeindruckt das vollständige Liniendiagramm sicherlich, doch fällt die Interpretation auf den ersten Blick noch recht schwer. Erkennbar ist aber zumindest, dass die Dominanz eines Faktors, wie ihn die Faktorenanalyse identifiziert hat, tatsächlich gegeben scheint. Diese Feststellung lässt sich daraus folgern, dass das Liniendiagramm bis auf den Mittelteil sehr linear verläuft. Zudem wird auch die Hierarchie der Sendungen bestätigt. Die durchschnittlich am besten benotete Sendung (*Die Simpsons*) landet tatsächlich ganz oben, während die am schlechtesten bewertete Sendung (*Big Brother*) sich ganz unten wiederfindet. Selbst die fünf Faktoren des Fernsehkonsums lassen sich in dem Liniendiagramm wiederfinden, so dass das Ergebnis der Faktorenanalyse tatsächlich bestätigt wird. Diese Aussage lässt sich besser durch eine Zerlegung des Diagramms in vier Abschnitte belegen:

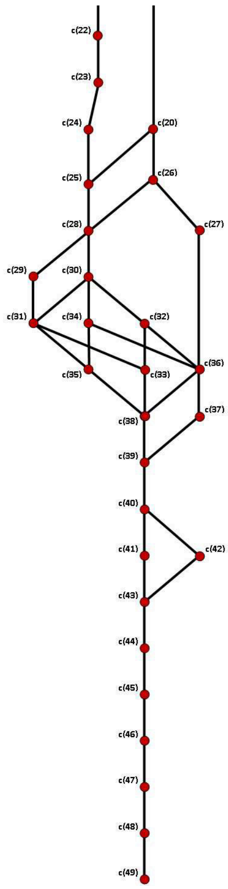


Abbildung 6.19 Ausschnitt 1 aus dem Liniendiagramm (latente Faktoren des Fernsehkonsums)

In diesem, von unten gesehen, ersten Block befinden sich nahezu alle Sendungen des ersten und zweiten Faktors, *gemischte Unterhaltung* und *Wettkampf und Bildung*. Dementsprechend stehen diese Sendungen in der Hierarchie weiter unten. Möglicherweise lässt sich das Ergebnis auch mit der zunehmenden Bedeutung anderer Medien, wie dem Internet und den Computerspielen erklären, die einen Teil der Unterhaltung Jugendlicher übernommen haben.

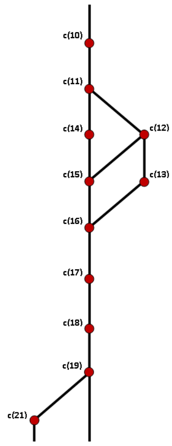


Abbildung 6.20 Ausschnitt 2 aus dem Liniendiagramm (latente Faktoren des Fernsehkonsums)

Der nachfolgende Abschnitt des Liniendiagramms beinhaltet die *niveauarmen Fernsehsendungen*. Auch diese liegen alle eng beieinander und größtenteils auf einer linearen Verbindung.

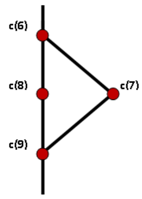


Abbildung 6.21 Ausschnitt 3 aus dem Liniendiagramm (latente Faktoren des Fernsehkonsums)

Auch die Sendungen zum Faktor *musikalische Unterhaltung* befinden sich in einer näheren Umgebung des Liniendiagramms. Interessant ist dabei, dass diese scheinbar sehr beliebt sind, aber es kaum noch „gutes" Musikfernsehen gibt, was durch die geringe Anzahl an Musiksendungen, die es in die finale Liste geschafft haben, deutlich wird. Vielleicht sollten sich die Fernsehsender gerade in diesem Bereich wieder etwas mehr einfallen lassen. Potenzielle Zuschauer scheinen vorhanden zu sein.

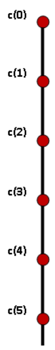


Abbildung 6.22 Ausschnitt 4 aus dem Liniendiagramm (latente Faktoren des Fernsehkonsums)

Abschließend bleibt zu bestätigen, dass auch die Sendungen des Faktors *humorvolle Unterhaltung* direkt beieinander und auf einer vollkommen linearen Verbindung liegen. Zudem sind die entsprechenden Sendungen ziemlich beliebt.

Die Auswertung der latenten Faktoren des Fernsehkonsums ist also mit Hilfe zweier Verfahren, deren einziges Kriterium *„Angepasstheit"* ist, durchgeführt worden. Da beide Vorgehensweisen zu einem sehr ähnlichen Ergebnis führen, kann die Methodenunabhängigkeit der gewonnenen latenten Faktoren angenommen werden. Damit wird zudem der Erfolg der 0-1-Daten bestätigt.

Abschließend sollen noch drei kleine Anmerkungen bezüglich der Philosophie der Datenerhebung und der Mathematik gestattet sein.

> (i) Zum einen ist die Faktorenanalyse auch für alle 80 Fernsehsendungen, nachdem zu diesem Datensatz ebenfalls die *missing values* nach dem beschriebenen Verfahren geschätzt wurden, durchgeführt worden. Dabei ergibt sich ein ähnliches Ergebnis mit allerdings acht zu unterscheidenden Faktoren. Der Faktor *humorvolle Unterhaltung* kann dabei in die Bereiche *Zeichentrick* und *Sitcom*, der Faktor *Wettkampf und Bildung* in eben diese Sektionen und der Faktor *niveauarme Unterhaltung* in die Gruppen *Doku-Soaps* und *Castingshows* unterteilt werden. Da diese Aufsplitterung allerdings sehr kleinschrittig ist, soll im weiteren Verlauf mit den fünf Faktoren der 53 Fernsehsendungen geforscht werden. Dennoch bestätigt diese Auswertung sowohl die Qualität der *missing values*-Schätzung als auch die Stabilität der extrahierten Faktoren.
>
> (ii) Zum anderen ist die Faktorenanalyse für die 53 Sendungen ebenfalls ohne Konvertierung der ordinalen Rohdaten in die 0-1-Daten durchgeführt worden. Auf Basis dieser Rohdaten ergibt sich ein unbrauchbares Ergebnis von zehn (!) inhaltlich nicht zu interpretierenden Faktoren. Somit wird das Antwortverhalten der Probanden tatsächlich falsch interpretiert, wenn das vorgestellte Verfahren keine Anwendung findet. Generell kann davon ausgegangen werden, dass in der Praxis solche Fehlinterpretationen nahezu immer vorkommen, mindestens aber dann, wenn über eines der Items (hier: Fernsehsendungen) eine relativ hohe Genauigkeit in der Übereinstimmung besteht bzw. keine extremen Abweichungen existieren. Dementsprechend verändert sich die Korrelationsmatrix bei affinen Transformationen nicht, wobei die einzelne Zufallsvariable dennoch in eine völlig andere Gruppe fallen würde. Das Problem der Fehlinterpretation, benutzt man weiterhin die unveränderten Rohdaten, tritt also nur dann nicht auf, wenn die Probanden bei jedem einzelnen Item in ihrem Abstimmungsverhalten extrem heterogen sind, wobei dann die Faktorenanalyse überhaupt kein brauchbares Er-

gebnis liefern dürfte. Folglich ist die Dichotomisierung der ordinalen Daten dringend notwendig und jedem Anwender ähnlicher empirischer Untersuchungen zu empfehlen.

(iii) Dass man sich um komplexere datenanalytische Methoden bemühen muss, liegt zudem daran, dass durch die Befragung der einzelnen Probanden das spezielle Datum von jedem einzelnen Probanden erhoben wird. Dieses spezielle Datum hat nun natürlich auch eine individuelle Datenstruktur, die vom Probanden ausgeht. Schließlich weiß nur der Proband selbst, warum er welcher Fernsehsendung welche Beurteilung zugewiesen hat. Wenn jetzt ein Verfahren der Datenanalyse genutzt wird, dann erhalten die Daten ihre Struktur allerdings nicht mehr durch den Probanden, sondern sie erhalten die Struktur durch das datenanalytische Verfahren. Es besteht die Möglichkeit, dass das Verfahren den Daten eine Struktur gibt, die an der eigentlichen Struktur, die der Proband beabsichtigt hat, vorbeiläuft. Das bedeutet für die Ratingskala in dem vorliegenden Fall, dass eine vom Probanden verteilte „2" als doppelt so gut interpretiert wird, wie eine vom Probanden vergebene „4". Dies ist aber keine Information, die den Daten entnommen werden darf, denn das Einzige, was man mit Sicherheit sagen kann, ist, dass der Proband die Fernsehsendung mit der „2" besser findet, als die Sendung mit der „4". Dementsprechend sollte kein Verfahren benutzt werden, dass den Daten eine Struktur gibt, die beim besten Willen nicht hineininterpretiert werden darf. Ganz im Gegenteil: Man muss sich auf das konzentrieren, was definitiv und tatsächlich ablesbar ist.

6.3.4 Verteilung der Probanden auf die fünf latenten Faktoren des Fernsehkonsums

Um die fünf identifizierten latenten Faktoren des Fernsehkonsums mit Leben zu füllen, gilt es, die 232 Probanden entsprechend ihres Antwortverhaltens in fünf Gruppen ähnlicher Fernsehgewohnheiten einzuteilen, damit für jeden Probanden ein konkreter Faktor für den Fernsehgebrauch angegeben werden kann. Dies geschieht mittels einer hierarchischen Clusteranalyse (Methode: Complete-Linkage, Proximitätsmaß: quadrierte Euklidische Distanz). Man vergleiche dazu die Ausführungen aus dem mathematischen Abschnitt „Verteilung der Probanden auf die latenten Faktoren". Das Verfahren liefert das folgende in Tabelle 6.7 illustrierte Ergebnis:

Tabelle 6.7

Bevorzugte Fernsehsendungen

	gemischte Unterhaltung	Wettkampf und Bildung	niveauarme Unterhaltung	humorvolle Unterhaltung	musikalische Unterhaltung
Probanden	46,6 %	28,4 %	16,8 %	3,4 %	4,7 %

Ein Großteil der Jugendlichen favorisiert kein konkretes Fernsehprogramm, sondern variiert je nach Lust und Laune zwischen anspruchsvollen und niveauarmen Fernsehsendungen: **46,6 %** der Probanden gehören dem dominanten ersten Faktor *gemischte Unterhaltung* an. Während **16,8 %** der befragten Jugendlichen vorwiegend *niveauarmes* Fernsehen (Faktor 3) konsumieren, schauen immerhin **28,4 %** am liebsten Sendungen, die dem Faktor *Wettkampf und Bildung* zuzuordnen sind. Der *humorvollen Unterhaltung* (Faktor 4) und der *musikalischen Unterhaltung* (Faktor 5) gehören dagegen kaum Probanden an (nur **3,4** und **4,7** %), so dass für diese beiden Gruppen keine statistisch aussagekräften Ergebnisse zu erwarten sind.

Einen detaillierten Einblick in das Zustandekommen dieser Auswertungsergebnisse bietet die Analyse der geschlechts- und schulformspezifischen Verteilungen der Probanden auf die fünf Faktoren in den beiden nachfolgenden Tabellen:

Tabelle 6.8

Bevorzugte Fernsehsendungen und Geschlecht

	gemischte Unterhaltung	Wettkampf und Bildung	niveauarme Unterhaltung	humorvolle Unterhaltung	musikalische Unterhaltung
männlich	62,7 %	20,6 %	7,1 %	2,4 %	7,1 %
weiblich	27,4 %	37,7 %	28,3 %	4,7 %	1,9 %

Tabelle 6.8 präsentiert zunächst das bevorzugte Fernsehprogramm in Abhängigkeit vom Geschlecht der Probanden. Interessanterweise favorisieren die meisten Mädchen Fernsehsendungen des Faktors *Wettkampf und Bildung*. Dieser positive Eindruck wird allerdings durch das ebenfalls überdurchschnittlich große Vergnügen an *niveauarmer Unterhaltung* direkt relativiert. Gerade die Mädchen schauen gerne Scripted-Reality-Formate und Castingshows! Die

gemischte Unterhaltung folgt letztlich auf dem dritten Platz und beinhaltet weniger als ein Drittel der weiblichen Untersuchungsteilnehmer.

Den Jungen fällt es offenbar relativ schwer, sich für eine konkrete Form an Fernsehsendungen zu begeistern: Überdurchschnittlich viele männliche Probanden werden beim Fernsehen am liebsten *gemischt* unterhalten. Wenngleich dieser erste Faktor durchaus einige anspruchslose Sendungen enthält, ist positiv hervorzuheben, dass die Jungen kaum *niveauarme Unterhaltung* bevorzugen. Ganz im Gegenteil: Dreimal so viele Jungen schauen lieber Formate des latenten Faktors *Wettkampf und Bildung*.

Die beiden letzten Faktoren besitzen statistisch zwar keine bedeutende Aussagekraft. Nichtsdestotrotz sollen zwei überraschende Anmerkungen nicht verschwiegen werden: Auf niedrigem Niveau schauen doppelt so viele Mädchen bevorzugt *Comedy*-Sendungen, während fast viermal mehr Jungen die *musikalische Unterhaltung* favorisieren. Hier hätte man sicherlich mit einem umgekehrten Ergebnis gerechnet.

Tabelle 6.9

Bevorzugte Fernsehsendungen und Schulform

	gemischte Unterhaltung	Wettkampf und Bildung	niveauarme Unterhaltung	humorvolle Unterhaltung	musikalische Unterhaltung
Gymnasium	61,1 %	27,8 %	10,0 %	0,0 %	1,1 %
Gesamtschule	39,4 %	35,6 %	21,2 %	1,9 %	1,9 %
Hauptschule	31,6 %	10,5 %	21,1 %	15,8 %	21,1 %

Ein abschließender Blick gebührt dem schulformspezifischen Fernsehprogramm, welches in Tabelle 6.9 dargestellt wird. Das Auswertungsergebnis der Gymnasiasten gleicht dem Auswertungsergebnis der Jungen in erheblichem Maße: Fast zwei Drittel bevorzugen die *gemischte Unterhaltung* und schwanken dabei regelmäßig im Anspruch der verschiedenen Fernsehsendungen. Pure *niveauarme* Fernsehsendungen werden lediglich von wenigen gymnasialen Probanden besonders gerne geschaut. Während *humorvolle* und *musikalische Unterhaltung* so gut wie gar nicht präferiert werden, können *Wettkampfs- und Bildungsformate* immerhin bei knapp einem Drittel der Probanden punkten.

Das bevorzugte Fernsehprogramm der Schülerinnen und Schüler der Gesamtschule gleicht der gymnasialen Verteilungen nur insofern, als dass die letzten beiden Faktoren keine nennenswerte Rolle spielen. Die Faktoren *gemischte Unterhaltung*, leicht unterhalb des Durchschnitts, und *Wettkampf und Bildung*, knapp oberhalb des Mittelwerts, liegen recht dicht beieinander auf den ersten beiden Plätzen. Allerdings wird auch die *niveauarme Unterhaltung* von über einem Fünftel der Probanden besonders gerne und oft geschaut.

Die Schülerinnen und Schüler der Hauptschule konsumieren sehr selten Formate, in denen die Vermittlung von *Bildung* oder das Verfolgen von *Wettkämpfen* im Mittelpunkt stehen. Nur jeder zehnte Proband ist für derartige Fernsehsendungen zu begeistern. Die Hauptschülerinnen und -schüler bevorzugen stattdessen die *gemischte Unterhaltung*, *niveauarme* Reality-Dokumentationen und *musikalische* Chartshows. Auch die *humorvollen* Comedy-Sendungen stehen bei diesen Jugendlichen vergleichsweise hoch im Kurs.

6.3.5 Zusammenfassung „latente Faktoren des Fernsehkonsums"

Der zweite zentrale Aspekt des Forschungsprojekts besteht in der Ermittlung der latenten Faktoren des Fernsehkonsums sowie der Verteilung der Jugendlichen auf eben diese. Dazu sind in einem aufwendigen Verfahren diejenigen Sendungen identifiziert worden, die sich die Probanden tatsächlich ansehen. Die 80 meistgesehenen Fernsehsendungen werden den Schülerinnen und Schülern schließlich mit der Bitte, diese in ihre individuelle Prioritätsordnung zu überführen vorgelegt. Der daraus resultierende Datensatz wird aufgrund einer Vielzahl fehlender Werte auf 53 Sendungen reduziert. Die übrig gebliebenen fehlenden Werte werden mit Hilfe eines eigenen Verfahrens geschätzt. Auf den nun vollständigen und in 0-1-Daten transformierten Datensatz kann die Hauptkomponentenmethode mit anschließender Varimax-Rotation angewendet werden. Dabei zeigt sich, dass fünf latente Faktoren den Fernsehkonsum von Jugendlichen beschreiben. Diese sind im Einzelnen *gemischte Unterhaltung*, *Wettkampf und Bildung*, *niveauarme Unterhaltung*, *humorvolle Unterhaltung* und *musikalische Unterhaltung*. Ein an die formale Begriffsanalyse angelehntes Verfahren bestätigt das Ergebnis, welches somit methodenunabhängig ist.

Die Verteilung der Probanden auf die gewonnenen Faktoren belegt, dass ein Großteil der Jugendlichen *gemischte Unterhaltungssendungen* bevorzugt. Auf den Plätzen zwei und drei folgen die Faktoren *Wettkampf und Bildung* und *niveauarme Unterhaltung*. Die letzten beiden Faktoren spielen statistisch keine relevante Rolle, wenn überhaupt an der Hauptschule.

6.4 Die latenten Faktoren des Freizeitverhaltens

Den Abschluss der fragebogenbasierten Datenerhebung bildet die Untersuchung der sonstigen Freizeitbeschäftigungen der jugendlichen Probanden. Das Ziel ist es, die Art der Freizeitbeschäftigung jedes einzelnen Probanden individuell zu bestimmen. Damit dies gelingen kann, müssen, ähnlich wie bei der Erhebung der latenten Faktoren des Fernsehkonsums, zunächst diejenigen Freizeitbeschäftigungen bestimmt werden, denen Jugendliche der vorliegenden Altersgruppe tatsächlich nachgehen. Im Anschluss daran ist es dann erneut die Aufgabe der Probanden, die in Schritt 1 gewonnenen Aktivitäten in ihre jeweilige individuelle Prioritätsordnung zu überführen. Abschließend können wiederum mit der für ordinalskalierte Daten optimierten Faktorenanalyse die latenten Faktoren der Freizeitbeschäftigung bestimmt werden. Auf Basis dieses Ergebnisses wird es dann möglich sein, jedem Probanden sein jeweiliges Freizeitverhalten zuzuordnen.

6.4.1 Datenerhebung

Um eine Liste sämtlicher Freizeitbeschäftigungen zu erhalten, werden die Probanden in einem ersten Fragebogen darum gebeten, ihre Hobbys anzugeben. Der zugehörige Arbeitsauftrag lautet:

> *„Wahrscheinlich schaust du in deiner Freizeit nicht nur TV. Gebe hier bitte deine weiteren Hobbys an!"*

Nach diesem Schritt ist es möglich, die genannten Freizeitbeschäftigungen statistisch festzuhalten. Auch an dieser Stelle sollen der geneigten Leserin und dem geneigten Leser die zehn meist genannten Antworten nicht vorenthalten bleiben:

Tabelle 6.10

Die Top 10 der Freizeitaktivitäten

Aktivität	Stimmen
Freunde treffen	157
Comupterspiele/Konsolenspiele	102
Fußball spielen	94
Schwimmen	93
Internet	88
Shoppen	80
Musik hören	78
Fahrrad fahren	62
Chillen	45
Tanzen	38

Insgesamt haben die Schülerinnen und Schüler 111 verschiedene Hobbys aufgezählt. Dabei fällt allerdings schnell die stark sinkende Anzahl der Nennungen durch die Probanden auf. Bedenkt man beispielsweise, dass der immerhin zehntbeliebtesten Beschäftigung (*tanzen*) gerade einmal 16 % der Probanden nachgehen, überrascht es nicht, dass etwa 30 Hobbys von nur jeweils einem einzigen Jugendlichen genannt werden. Derartig selten vorkommende Angaben sind natürlich für die finale Liste derjenigen Freizeitaktivitäten, die die Schülerinnen und Schüler letztendlich ihrer Beliebtheit nach ordnen müssen, vollkommen ungeeignet. Praktischer erscheint es dagegen, analog zum Vorgehen der JIM-Studie, seltener genannte Antworten zu Gruppen ähnlicher Aktivitäten zusammenzufassen. Beispielsweise werden die Probandenantwort *Ausflüge mit Freunden* dem Komplex *Freunde treffen* zugeordnet, sämtliche sportliche Aktivitäten unter dem naheliegenden Motto *Sport* zusammengefasst oder das Spielen verschiedener Musikinstrumente als *Musik machen* bezeichnet. Auf diese Art und Weise gelingt es sowohl sämtliche Antworten zu berücksichtigen, als auch die unübersichtlich lange Liste auf ein vernünftiges Maß zu reduzieren.

Nachdem am Ende des ersten Schrittes letztlich eine Liste von 19 Freizeitbeschäftigungen erstellt werden kann, ist es nun abermals die Aufgabe der Schülerinnen und Schüler, diese Beschäftigungen der jeweiligen individuellen Beliebtheit nach zu sortieren. Da mit ähnlichen Problemen wie bei Schritt 2 zur Ermittlung der latenten Faktoren des Fernsehkonsums gerechnet werden muss, werden die Probanden gleichermaßen im Vorfeld der Datenerhebung einer Reflexionsphase unterzogen und gebeten, die verschiedenen Aktivitäten zu benoten, um als Grundlage der Faktorenanalyse zur Ermittlung der latenten Faktoren der Freizeitbeschäftigung den Daten erneut zumindest ordinale Signifikanz durch die Probanden unterstellen zu können.

6.4.2 Die Ermittlung der latenten Faktoren des Freizeitverhaltens

Die Ermittlung der latenten Faktoren des Freizeitverhaltens auf Basis des vollständigen Datensatzes verläuft identisch, wie die bereits vorgestellte Ermittlung der latenten Faktoren des Fernsehkonsums. Da das Verfahren als bekannt vorausgesetzt wird, sollen an dieser Stelle lediglich die Ergebnisse vorgestellt werden.

Die Hauptkomponentenmethode mit anschließender Varimax-Rotation ermittelt dabei zwei latenten Faktoren, die das Freizeitverhalten der Rezipienten beschreiben:

(i) *aktiv/kreativ:* Der erste Faktor beinhaltet vor allem kreative („Schreiben", „Bücher lesen", „Malen, Basteln", „Selbst Musik machen", „Kochen", „Fotos machen") oder aktive („Haustiere", „Familienunternehmungen", „Sport") Hobbys. Selbst die diskussionswürdigen „Computerspiele" erfordern, verglichen mit dem

Fernsehen, eine aktive Auseinandersetzung. Daher kann der Faktor als **aktiv/kreative Freizeitbeschäftigung** interpretiert werden.

(ii) *passiv/kognitiv anspruchslos:* Der zweite Faktor umfasst dagegen eher passive Beschäftigungen („Handy", „Musik hören", „Internet", „Filme schauen", „Ausruhen/Nichts tun", „Freunde treffen") bzw. kognitiv weniger anspruchsvolle Aktivitäten („Shoppen", „Party"). Basierend auf diesen Feststellungen soll der zweite Faktor als **passive bzw. kognitiv anspruchslose Freizeitbeschäftigung** bezeichnet werden.

In der nachfolgenden Tabelle 6.11 wird die detaillierte Verteilung der Aktivitäten auf die zwei extrahierten Faktoren dargestellt. Die Faktorenanalyse führt dementsprechend erneut zu einem sehr zufriedenstellenden Ergebnis.

Tabelle 6.11

Latente Faktoren des Freizeitverhaltens

aktiv/kreativ	passiv/kognitiv anspruchslos
Schreiben Bücher lesen Computer (offline) Malen, Basteln Selbst Musik machen Kochen Haustiere Familienunternehmungen Fotos machen Computerspiele Sport	Handy Musik hören Internet Filme schauen Ausruhen, Nichts tun Mit Freunden treffen Partys Shoppen

6.4.3 Aktive und passive Freizeitbeschäftigung

Nachdem nun zunächst geklärt worden ist, dass sich das Freizeitverhalten der Probanden in *aktiv/kreative* und *passiv/kognitiv anspruchslose* Aktivitäten unterscheiden lässt, gilt es, die Probanden auf die beiden Freizeittypen zu verteilen. Dies geschieht erneut durch Anwendung des clusteranalytischen Verfahrens aus dem Abschnitt „Verteilung der Probanden auf die latenten Faktoren" aus dem mathematischen Kapitel (Methode: Complete-Linkage, Proximitätsmaß: quadrierte Euklidische Distanz).

Dazu betrachte man Tabelle 6.12:

Tabelle 6.12

Bevorzugtes Freizeitverhalten

	aktiv/kreativ	passiv
Probanden	43,1 %	56,9 %

Insgesamt zeigt das Ergebnis der Studie, dass ein Großteil der Jugendlichen eine *passive* Freizeitgestaltung präferiert: Während lediglich **43,1 %** der Probanden *aktive* und *kreative Freizeitaktivitäten* bevorzugen, wählen **56,9 %** eher *passive* und *kognitiv weniger anspruchsvolle Beschäftigungen* in ihrer Freizeit aus. Bei den aktiven Freizeitbeschäftigungen erfreuen sich vor allem die Hobbys „Sport" und „Computerspiele" großer Beliebtheit, während bei den passiven Freizeitaktivitäten vor allem „chillen mit Freunden" besonders gute Bewertungen erhalten hat.

Detaillierte und teilweise überraschende Erkenntnisse liefert dagegen die Betrachtung der verschiedenen Geschlechter und Schulformen in den folgenden Tabellen:

Tabelle 6.13

Freizeitverhalten und Geschlecht

	aktiv/kreativ	passiv
männlich	50,0 %	50,0 %
weiblich	34,9 %	65,1 %

Die Tabelle 6.13 präsentiert dabei den Anteil der männlichen und weiblichen Probanden hinsichtlich ihrer Freizeitbeschäftigungen. Während das Verhältnis zwischen *passiven* und *aktiven* männlichen Jugendlichen in dieser Studie exakt ausgeglichen ist, überrascht es auf den ersten Blick durchaus, dass gerade bei den weiblichen Probanden *passive* Formen der Freizeitgestaltung mit **65,1 %** klar im Vorteil sind. Besonders die Hobbys „Handy", „Internet", „Shoppen" und „Partys" konnten dabei als *typisch* weiblich identifiziert werden und legen den Grundstein für dieses Ergebnis.

Tabelle 6.14

Freizeitverhalten und Schulform

	aktiv/kreativ	passiv
Gymnasium	61,1 %	38,9 %
Gesamtschule	25,0 %	75,0 %
Hauptschule	50,0 %	50,0 %

In der Tabelle 6.14 werden nun abschließend die schulformabhängigen Forschungsergebnisse dargestellt. Dabei sind vor allem die Schülerinnen und Schüler des Gymnasiums darauf bedacht, in der schulfreien Zeit möglichst vielen *aktiven* und *kreativen* Aktivitäten nachzugehen. Über **60 %** der Gymnasiasten favorisieren ein solches Freizeitverhalten. Als *typisch gymnasiale* Hobbys gelten dabei neben den generell hoch eingestuften Aktivitäten „Sport" und „Computerspiele" vor allem „Schreiben", „Bücher", „Malen/Basteln" sowie „Musik machen", die allesamt überdurchschnittlich gut bewertet wurden. Ein komplett anderes Bild zeigt dagegen die Freizeitbeschäftigung der Gesamtschülerinnen und -schüler. Genau **75 %** dieser Probanden beschäftigen sich in ihrer Freizeit vorzugsweise *passiv*. Vor allem die *kreativen* Hobbys kommen bei der Bewertung besonders schlecht weg. Dieser eklatante Unterschied ist durchaus überraschend, ebenso wie das völlig ausgeglichene Verhältnis bei den Probanden, die die Hauptschule besuchen. Während hier vor allem „Partys", „Freunde treffen", „Ausruhen, Nichts tun" und „Internet" bevorzugte *passive* Freizeitbeschäftigungen sind, werden die *aktiven* und *kreativen* Aktivitäten hauptsächlich durch „Sport", „Computerspiele" und „(Digitale-) Fotos machen" repräsentiert.

6.4.4 Zusammenfassung „latente Faktoren des Freizeitverhaltens"

Die abschließende Untersuchungseinheit beschäftigt sich mit dem fernsehfreien Freizeitverhalten der Probanden, mit dem Ziel die zahlreichen Hobbys der Schülerinnen und Schüler in übersichtlichen Faktoren zusammenzufassen. Aus diesem Grund wird zunächst durch eine Umfrage eine Liste all derer Aktivitäten aufgestellt, denen die Jugendlichen tatsächlich nachgehen. Diese Freizeitbeschäftigungen können in Anlehnung an die JIM-Studie zu 19 unterschiedlichen Hobbys zusammengefasst werden. Jeder Proband überführt letztlich die aufgelisteten Aktivitäten in die eigene Prioritätsordnung. Auf den erhaltenen und in 0-1-Daten transformierten Datensatz kann nun die Hauptkomponentenmethode mit anschließender Varimax-Rotation angewendet werden. Dabei kristallisieren sich zwei („*kreativ/aktiv*" und „*pas-*

siv/kognitiv anspruchslos") das Freizeitverhalten von Jugendlichen beschreibenende latente Faktoren heraus.

Bei der Verteilung der Probanden auf die zwei Typen der Freizeitgestaltung wird gezeigt, dass die Mehrzahl der Probanden *passive* Freizeitbeschäftigungen favorisiert. Vor allem Mädchen und GesamtschülerInnen fallen hier auf.

Das Kapitel abschließend zusammenfassend, können die folgenden zentralen Ergebnisse präsentiert werden:

a) Die Probanden können bezüglich ihrer täglichen Fernsehdauer in vier Gruppen (*Wenigseher*, *Mittelseher*, *Vielseher*, *Exzessivseher*) eingeteilt werden.

b) Es gibt fünf, im inhaltlichen Anspruch stark schwankende, latente Faktoren für den Fernsehkonsum von Jugendlichen (*gemischte Unterhaltung*, *Wettkampf und Bildung*, *niveauarme Unterhaltung*, *humorvolle Unterhaltung*, *musikalische Unterhaltung*).

c) Zwei latente Faktoren (*kreativ/aktiv*, *passiv/kognitiv anspruchslos*) beschreiben das Freizeitverhalten der Schülerinnen und Schüler.

7. Möglichkeiten der Leistungsmessung - eine mathematisch-methodische Untersuchung

7.1 Leistungsmessung in der didaktischen Literatur

Nachdem die Datenerhebung des Fernseh- und Freizeitverhaltens als abgeschlossen gilt, ist es von elementarer Bedeutung die Mathematikleistungen der Probanden zu bestimmen, um im weiteren Verlauf Zusammenhänge untersuchen zu können. Bezüglich der *Leistungsmessung* existiert in der didaktischen Literatur eine lebendige Diskussion, wobei im Folgenden lediglich grob auf subjektiv ausgewählte Aspekte eingegangen werden kann, da es im Rahmen dieser Arbeit nicht möglich ist, die ganze Literatur zu reflektieren.

In einem Punkt sind sich die meisten Experten einig. Die Zensuren in der Schule sind kein geeignetes Messinstrument für die tatsächlich erbrachte Leistung der Schülerinnen und Schüler. Es wird beispielsweise kritisiert, dass die Notenstufen durch die Kultusministerkonferenz (KMK) äußerst unscharf definiert sind (vgl. Ziegenspeck 1999) und soziale Bezugsnormen, sowie nicht-fachlich relevantes Schülerverhalten ebenfalls einen nicht zu unterschätzenden Einfluss auf die Notenvergabe besitzt (vgl. Tillmann; Vollstädt 1999, S. 42 ff). Zahlreiche empirische Studien von Starch und Elliott (1977), Weiss (1989), Thiel und Valtin (2002), Eells und Dicker (1977) sowie Ingenkamp (2005) weisen zudem schon seit den 1970er Jahren auf die beschränkte messtheoretische Qualität der Notengebung in der Schule hin. Demnach genügen Zensuren den zentralen psychometrischen Gütekriterien *Objektivität, Reliabilität* und *Validität* nur unzureichend. Konkret für den Fachbereich der Mathematik gelang es in einer groß angelegten Studie nachzuweisen,

> *„dass sich in den analysierten (Zeugnis-) Noten auch bei Kontrolle des über Tests erfassten Leistungsstands verschiedene weitere Schülermerkmale (wie das Fachinteresse, die Anstrengung des Schülers) widerspiegeln, dass sich aber die Regressionsgewichte der Testleistung, von 'leistungsnahen' Merkmalen (z.B. Fachinteresse) und 'Hintergrundmerkmalen' (z.B. Geschlecht) signifikant und erheblich in Abhängigkeit von der Klassenzugehörigkeit unterscheiden"* (Hochweber 2010, S. 13).

So einheitlich die Meinung der Experten bezüglich der Notengebung auch sein mag, so unterschiedlich sind die geforderten Konsequenzen. Dabei lassen sich laut Vollstädt (2005, S. 15) zwei grundsätzliche Ansätze in der einschlägigen Literatur unterscheiden:

- Vertreter der ersten Richtung orientieren sich an der reformpädagogischen Diskussion der zwanziger Jahre, als ein Leistungsbegriff entwickelt wurde, der die Gesamtpersönlichkeit der Schüler in den Blick nahm, so dass die Schule zu einem Ort der Entfaltung individueller Entwicklungsmöglichkeiten werden sollte (vgl. ebd., S. 12). Das Ziel entsprechender Didaktiker ist daher, die Bedeutung von Zensuren weitestgehend einzuschränken und Noten durch Lernberichte zu ersetzten (vgl. ebd., S. 15). Gegner dieser reformpädagogischen Form der Leistungsmessung befürchten allerdings, dass Lernberichte die Schülerinnen und Schüler nicht ausreichend auf das Arbeiten und Leben in einer „Leistungsgesellschaft" vorbereiten würden (vgl. ebd., S. 12).
- Die Vertreter des zweiten Ansatzes zweifeln die Aussagekraft von Zensuren weniger aus ideologischen Gesichtspunkten an, da sie einen wissens- und stoffdominierten Leistungsbegriff durchaus befürworten. Vielmehr basiert die Abneigung dieser Didaktiker auf den Ergebnissen empirischer Untersuchungen, welche die Notengebung in Frage stellen. Entsprechende Vertreter suchen daher

  > *„nach Möglichkeiten und wissenschaftlich legitimierten Modellen zur Verbesserung der angewendeten 'Messinstrumentarien'. Sie suchen nach geeigneten Testverfahren und standardisierten Methoden zur Leistungsmessung, die vor allem eine höhere 'Objektivität' der Bewertung anstreben"* (ebd., S. 15).

Auch Brügelmann (2002) hält fest, dass es sich der Leistungsmessung in der Schule um ein komplexes Problem handelt, auch weil verschiedene Perspektiven (Bildungsforscher, Schulaufsicht, Lehrer, Eltern, Arbeitgeber, ...) berücksichtigt werden müssen. Dabei sollten drei Ebenen unterschieden werden:

> - *„die politische: Wer bestimmt die Kriterien, wer kontrolliert die Umsetzung - zentrale oder dezentrale Entscheidungsträger?*
> - *die forschungsmethodische: Welches Instrumentarium ist das aussagekräftigste und verlässlichste - die standardisierte Messung oder das personengebundene Urteil?*
> - *die reformstrategische: Auf welchen Wegen ist Schulentwicklung am ehesten in Gang zu bringen - durch interne oder durch eine externe Evaluation?"* (Brügelmann 2002, S. 41)

Als bekannte standardisierte Messinstrumentarien werden u. a. nationale (z.B. Bildungstest des IQB) und internationale (z.B. PISA-Studie) Bildungsstudien angesehen. Das *Programm zur internationalen Schülerbewertung* (PISA) misst beispielsweise die Leistungen der Schülerinnen und Schüler mit Hilfe probabilistischer Testmodelle. Die Tests fragen allerdings kein konkretes Schulwissen ab, sondern unterscheiden *„in allen Inhaltsbereichen zwischen Konzeption, Prozessen und Situationen beziehungsweise Kontexten"* und orientieren sich *„an einer Vorstellung von lebenslangem Lernen und betonen das Verstehen und die flexible, situationsgerechte Anwendung des Wissens"*(Deutsches PISA-Konsortium 2005, S. 32). Gerade diese Art der Leistungsmessung ist für deutsche Schülerinnen und Schüler unbekanntes Neuland gewesen (vgl. ebd., S. 32).

Die Rückmeldung über die gemessene Leistung erfolgt bei PISA, ähnlich wie bei Klassenarbeiten in der Schule, über den von den Schülerinnen und Schülern erreichten Punktwert. Im Unterschied zu Schulnoten erfolgt die Ergebnisdarstellung mitsamt ihrer weitreichenden Deutung hinsichtlich der Fähigkeiten, Stärken und Schwächen sowie Bedingungsfaktoren über fünf *Kompetenzstufen*, welche die erreichten Punkte anschaulich darstellen und kommunizieren sollen (vgl. Meyerhöfer o.J.).

Kritiker befürchten, dass die Vieldimensionalität der (Mathematik-) Fertigkeiten zugunsten einer konkreten Zuordnung jeder Schülerin und jedes Schülers zu einer der fünf abstrakten Kompetenzstufen geopfert wird. Die Ergebniswiedergabe bekommt damit, ähnlich wie bei Schulnoten, einen eindimensionalen Charakter (vgl. Deutsches PISA-Konsortium 2005, S. 53 f). Zudem wird kritisiert, dass keine Lernfortschritte gemessen werden und es daher unklar bleibt, ob die Erfolge und Misserfolge einzig und allein der Schule zuzuschreiben sind (vgl. Gaugner 2010, S. 10 zitiert nach Fuhrmann; Beckmann-Dierkes 2011, S. 9). Außerdem müssen sich die Entwickler der PISA-Studie den Vorwurf gefallen lassen, strukturelle Verschiedenheiten von Bildungssystemen und deren ebenso unterschiedlichen Voraussetzungen zu ignorieren (vgl. Fuhrmann; Beckmann-Dierkes 2011, S. 9). Befürworter wie Manfred Prenzel (2008) sind dahingegen davon überzeugt, dass sich das deutsche Schulsystem nach den (unter-) durchschnittlichen Ergebnissen der ersten PISA-Studien bereits deutlich verbessert hat und sich das erhöhte Engagement der Lehrkräfte in der Schul- und Unterrichtsentwicklung widerspiegelt.

Um die Wirksamkeit des (Mathematik-) Unterrichts in Deutschland zu verbessern, hat die KMK zur Weiterentwicklung der Bildungsqualität gemeinsame Bildungsstandards eingeführt. Diese Bildungsstandards sollen einen kompetenzorientierten Unterricht ermöglichen und als Grundlage für die Überprüfung der Schülerleistungen dienen. Das Konzept der Bildungsstandards legt dazu fest, welche fachbezogenen Kompetenzen (d. h. *„Wissen und Können in den jeweiligen Fächern zur Lösung von Problemen anzuwenden"* (Kultusministerkonferenzen 2012, S. 1)) die Schülerinnen und Schüler zu bestimmten Eckpunkten ihrer Schullaufbahn entwickelt haben sollen (vgl. ebd., S. 1).

Vollstädt (2005, S. 14) ist daher der Meinung, dass der veränderten Unterrichtskultur auch in der Leistungsmessung Rechnung getragen werden muss:

> *„Es muss ein in der Praxis der Schulen verwendbares Instrumentarium erarbeitet werden, das den Verlauf und die Ergebnisse problemlösender, handlungsorientierter, selbstbestimmter, eigenverantwortlicher und kooperativer Lernprozesse erfassen und bewerten kann. Dieses Instrumentarium verknüpft die Möglichkeiten der Selbstbeobachtung und Selbsteinschätzung mit den Formen der Fremdeinschätzung und unterstützt die Entwicklung einer Beobachtungs- und Beurteilungskompetenz auf Seiten der Lehrenden und Lernenden".*

Hinsichtlich der Leistungsmessung im Fachgebiet der Mathematik sind STERN und HARDY (2002, S. 168) durchaus zuversichtlich, dass eine sinnvolle Leistungsmessung möglich ist. Sie fordern dazu allerdings einen Kompromiss zwischen den an der Diskussion beteiligten Akteuren. Das Mindestmaß an Übereinstimmung sollte beinhalten, dass Problemlösen als zentrale mathematische Kompetenz angesehen wird und daher immer neue Aufgaben entwickelt werden müssen, die diesen Bereich testen. Zudem müsse die Förderung mathematischer Kompetenzen als langfristiger Prozess und Leistungsmessung als Indikator für die Unterrichtsqualität begriffen werden.

Aufgabe dieses Kapitels soll es nicht sein, didaktisch motiviert in diese Diskussion einzugreifen, weil sich die Auffassungen teilweise widersprechen und der Verfasser nicht irgendeiner Meinung einen Vorzug geben will. Dies entspräche lediglich einer subjektiven Auswahl und genügt somit keinen wissenschaftlichen Ansprüchen. Da es das finale Ziel dieser Arbeit ist, Zusammenhänge zwischen Fernsehverhalten und (im Unterricht gemessenen) Mathematikleistungen zu untersuchen, kann die Thematik allerdings auch nicht vollständig ausgeklammert werden. Daher soll ein methodisch anderer Versuch unternommen werden, in diese Diskussion einzugreifen, indem versucht wird, die Situation der Leistungsmessung mathematisch zu modellieren, um anschließend zu hinterfragen, ob diese Modelle Lösungen zulassen, welche (möglichen) Expertenmeinungen optimal angepasst sind.

Dazu sind prinzipiell zwei Anforderungen zu erfüllen:

a) Auswahl der geeigneten Kriterien, die zur Leistungsmessung herangezogen werden.
b) Realisierung der Leistungsmessung, wenn entsprechende Kriterien vorliegen.

Dieses Kapitel gliedert sich dementsprechend in drei Abschnitte. Die ersten beiden Teile widmen sich dabei der mathematisch-methodischen Analyse der zu bewältigenden Aufgaben. Abschließend wird ein Vorschlag zur konkreten Leistungsmessung in diesem Forschungsprojekt vorgestellt.

7.2 Kriterien zur Leistungsmessung

Zunächst soll die Auswahl der geeigneten Kriterien, die zur Leistungsmessung herangezogen werden, im Mittelpunkt der Modellierung stehen. Üblicherweise ist Modellierung in der Mathematik nie eindeutig. Dies bedeutet, dass ein Modell, welches die Situation der Leistungsmessung auf der Grundlage von Expertenmeinungen modelliert ebenfalls nicht eindeutig ist. Aus diesem Grund werden im Folgenden drei Modelle messtheoretischer Natur vorgestellt, die sich in kanonischer Weise aus den Arbeiten von HERDEN und PALLACK (2005), HERDEN (1999) und BECKE (1996) ableiten lassen. Die Modelle orientieren sich an den Kriterien *Entwicklung einer adäquaten Zielfunktion*, *Stabilität von Rangordnungen* und dem Begriff des *Konsenskriteriums.*

Das erste Modell ist sowohl mess- als auch wahrscheinlichkeitstheoretischer Natur. Dazu betrachten wir einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, bei dem die Urne Ω als die Menge der Kriterien zur Leistungsmessung, die den Experten zur Verfügung stehen und beurteilt werden sollen, interpretiert wird. Unter dieser Voraussetzung können die Experten durch reelle, messbare Zufallsvariablen modelliert werden. Da die Anzahl der Experten nicht von vornherein eingegrenzt werden soll, lassen wir beliebig viele abzählbare Experten zu. Die entsprechende Menge reeller Zufallsvariablen sei im Folgenden mit $\mathfrak{X}$ abgekürzt. Jede solche Zufallsvariable $X_i \in \mathfrak{X}$ repräsentiert dann einen Experten. Wähle man nun also ein beliebiges $X_i \in \mathfrak{X}$, so beschreibt $X_i(\omega)$ die Beurteilung des Kriteriums $\omega \in \Omega$ durch den Experten X_i. Im Regelfall stehen natürlich nicht unendlich viele Experten zur Verfügung, sondern lediglich eine vorgegebene endliche Teilmenge von Experten. Wir gehen dementsprechend davon aus, dass n Experten aus der Menge der Zufallsvariablen ausgewählt werden. Diese endliche Teilmenge sei durch $\left\{X_{i_1}, \ldots, X_{i_n}\right\}_{i \in I}$ gegeben.

Es ist nun das Ziel des ersten Modells eine möglichst optimale Rangfolge der Kriterien zu erstellen. Diese Rangfolge symbolisiert den optimalen Kompromiss der Experten. In der Regel erreicht man dies durch die Minimierung einer Zielfunktion $G : \mathfrak{X}^{n+1} \mapsto \mathbb{R}$. Dies bedeutet, dass eine Zufallsvariable $X \in \mathfrak{X}$ gesucht wird, für die die Zielfunktion $G : \mathfrak{X}^{n+1} \mapsto \mathbb{R}$ die Gleichung $G(X, X_{i_1}, \ldots, X_{i_n}) = \min_{X' \in \mathfrak{X}} G(X', X_{i_1}, \ldots, X_{i_n})$ erfüllt. X ist dann diejenige Zufallsvariable, die in bestmöglicher Weise die einzelnen Expertenmeinungen repräsentiert.

Nun gilt es zu untersuchen, inwieweit geeignete Zielfunktionen $G : \mathfrak{X}^{n+1} \mapsto \mathbb{R}$ existieren. Ein einfaches Beispiel ist die, in den Sozialwissenschaften weitverbreitete, Mittelwertfunktion. Im Fall der Leistungsmessung würde dies bedeuten, dass die Mittelwerte aller Einzelmeinungen gebildet werden, deren Werte für die einzelnen Kriterien dann die Rangfolge der Kriterien widerspiegelt. Scheinbar stehen also derartige Funktionen zur Verfügung. Es ist aber im Einzelfall zu überprüfen, ob diese Funktionen auch wirklich geeignet sind. Den Begriff des

„Geeignetseins" präzisieren wir im Folgenden in erster Annäherung durch den Begriff der *Skalenadäquatheit* von G. Das meint, dass das G mit dem Datenniveau der Zufallsvariablen $X_i \in \mathfrak{X}$ verträglich sein muss.

Die Frage nach der Skalenadäquatheit von G ist bereits in einem wesentlich allgemeineren Kontext unter dem Oberbegriff *Skalenadäquate Zielfunktionen* von HERDEN und PALLACK (2005) untersucht worden. Trotzdem verfolgen wir hier eine noch allgemeinere Idee, die in Zusammenarbeit mit meinem Doktorvater PROF. HERDEN entwickelt wurde.

Zur mathematischen Präzisierung des Begriffs der *Skalenadäquatheit* betrachte man die fest gewählte linear (total) geordnete Menge $(\mathbb{R}, \leq)$ der reellen Zahlen. Auf $(\mathbb{R}, \leq)$ betrachte man die σ-Algebra der Borelmengen. Für den auf diese Weise definierten Messraum $(\mathbb{R}, \mathcal{B})$ besteht dann die Gruppe $(Aut(\mathbb{R}, \leq), \circ)$ aller Ordnungsautomorphismen $T : (\mathbb{R}, \leq) \mapsto (\mathbb{R}, \leq)$ aus messbaren Abbildungen. Darüber hinaus sei eine beliebige unendliche, aber abzählbare Familie $\{(\Omega_i, \mathcal{A}_i)\}_{i \in I}$ von Messräumen gegeben (die Unendlichkeit dieser Familie ist notwendig, um zuzulassen, dass beliebig viele Beobachtungen (Messungen) vorliegen können). Auf jedem Messraum $(\Omega_i, \mathcal{A}_i)$ seien sowohl eine Menge $\mathfrak{X}_i$ von (beschränkten) messbaren Abbildungen $X_i : (\Omega_i, \mathcal{A}_i) \mapsto (\mathbb{R}, \mathcal{B})$ als auch eine bezüglich abzählbaren Vereinigungen und abzählbaren Durchschnitten abgeschlossene Teilmenge $\mathcal{N}_i$ von $\mathcal{A}_i$, die es erlaubt zwei messbare Abbildungen $Y : (\Omega_i, \mathcal{A}_i) \mapsto (\mathbb{R}, \mathcal{B})$ und $Z : (\Omega_i, \mathcal{A}_i) \mapsto (\mathbb{R}, \mathcal{B})$ als *fast gleich* anzusehen, wenn die Menge aller $\omega \in \Omega_i$, für die $Y(\omega) \neq Z(\omega)$ gilt, zu $\mathcal{N}_i$ gehört, ausgezeichnet. Des Weiteren sei unterstellt, dass jede Menge $\mathfrak{X}_i$ eine messbare Abbildung 0_i enthält, die *fast überall verschwindet*, für die also die Menge aller $\omega \in \Omega_i$, für die $0_i(\omega) \neq 0$ gilt, in $\mathcal{N}_i$ enthalten ist. Diese Sprechweise meint insbesondere, dass $\mathcal{N}_i$ der Menge der Nullmengen eines Wahrscheinlichkeitsraums $(\Omega, \mathcal{A}, P)$ entspricht. Für alle Mengen $\mathfrak{X}_i$ ist das Datenniveau der in $\mathfrak{X}_i$ enthaltenen messbaren Abbildungen durch eine feste Untergruppe $(\mathbf{T}_i, \circ)$ von $(Aut(\mathbb{R}, \leq), \circ)$ bestimmt, die dementsprechend die Eigenschaft hat, dass für alle $T \in \mathbf{T}_i$ die Inklusion $T(\mathfrak{X}_i) := \{T \circ X_i \mid X_i \in \mathfrak{X}_i\} \subset \mathfrak{X}_i$ erfüllt ist. Ist $\prod_{i \in I} \mathfrak{X}_i$ das Kartesische Produkt der Mengen $\mathfrak{X}_i$ und $\sum_{i \in I} \mathfrak{X}_i$ die Menge aller Tupel $(X_i)_{i \in I} \in \prod_{i \in I} \mathfrak{X}_i$, für die, bis auf endliche viele Komponenten, X_i fast überall mit 0_i übereinstimmt, dann werde im Folgenden jede Funktion $G : \sum_{i \in I} \mathfrak{X}_i \mapsto \mathbb{R}$ als *Zielfunktion* (*Gütekriterium*) bezeichnet. Sei $\left(\prod_{i \in I} \mathbf{T}_i, \circ \right)$ das direkte Produkt der Gruppen $(\mathbf{T}_i, \circ)$. Dann betrachte man die Menge aller Tupel $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$, die die Eigenschaft haben, dass es höchstens endlich viele $i \in I$ gibt, für die $T_i(0)$ von 0 verschieden ist. Die Menge derjenigen $i \in I$, für die $T_i(0)$ von 0 verschieden ist, sei mit $A(T)$ abgekürzt. Dann werden im Folgenden drei Stufen zulässiger Transformationen $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ differenziert:

Ein Tupel $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ erfüllt das Kriterium der *Zulässigkeit erster Ordnung* (ist von *erster Ordnung*), wenn alle Gruppen $(\mathbf{T}_i, \circ)$ identisch sind und darüber hinaus für alle $i \in A(T)$ und alle $j \in A(T)$ die Transformationen T_i und T_j übereinstimmen. Ein Tupel $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ erfüllt das Kriterium der *Zulässigkeit zweiter Ordnung* (ist von *zweiter Ordnung*), wenn alle Gruppen $(\mathbf{T}_i, \circ)$ übereinstimmen. In allen übrigen Fällen erfüllt ein Tupel $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ das Kriterium der *Zulässigkeit dritter Ordnung*, was kürzer meint, dass es von *dritter Ordnung* ist. Im Folgenden wird eine mindestens zweielementige Teilmenge $\mathcal{D}$ von $\sum_{i \in I} \mathfrak{X}_i$ als *zulässige Datenmenge* bezeichnet, wenn eine *fest vorgegebene* endliche Teilmenge $Z(\mathcal{D})$ von I existiert, so dass für alle Tupel $(X_i)_{i \in I} \in \mathcal{D}$ und alle $i \in I$ die Äquivalenz „$X_i = 0_i$ fast überall $\Leftrightarrow i \in Z(\mathcal{D})$" gilt.

Nun sei unterstellt, dass man an der Minimierung der Zielfunktion interessiert ist. Dies bedeutet keine Einschränkung, da das gegenteilige Interesse dem Interesse an der Minimierung der Zielfunktion dual gegenüber steht, also durch Betrachtung der linear (total) geordneten Menge $(\mathbb{R}, \geq)$ an Stelle von $(\mathbb{R}, \leq)$ auf die Minimierung der Zielfunktion zurückgeführt werden kann.

Sei daher für jedes Tupel $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ die Menge $T(\mathcal{D})$ durch $T(\mathcal{D}) := \left\{ (T_i(X_i))_{i \in I} \middle| (X_i)_{i \in I} \in \mathcal{D} \right\}$ definiert und seien darüber hinaus eine Zielfunktion $G : \sum_{i \in I} \mathfrak{X}_i \mapsto \mathbb{R}$ sowie eine natürliche Zahl $1 \leq k \leq 3$ fest vorgegeben, dann besteht die *k-te Invarianzgruppe* $(\mathbf{I}_k(G_{\mathcal{D}}), \circ)$ von G bezüglich $\mathcal{D}$ aus allen zulässigen Tupeln $T = (T_i)_{i \in I} \in \prod_{i \in I} \mathbf{T}_i$ k-ter Ordnung, die sowohl die Gleichung $A(T) = Z(\mathcal{D})$ als auch die Äquivalenz $G\left(\left(X_i\right)_{i\in I}\right) = \min_{\left(Y_i\right)_{i\in I}\in\mathcal{D}} G\left(\left(Y_i\right)_{i\in I}\right) \Leftrightarrow G\left(\left(T\left(X_i\right)\right)_{i\in I}\right) = \min_{\left(Z_i\right)_{i\in I}\in T(\mathcal{D})} G\left(\left(Z_i\right)_{i\in I}\right)$ erfüllen. Nach diesen Vorüberlegungen lässt sich der Begriff der *Skalenadäquatheit k-ter Ordnung* (kurz: *k-Skalenadäquatheit*) einer Zielfunktion $G : \sum_{i\in I}\mathfrak{X}_i \mapsto \mathbb{R}$ wie folgt präzise fassen: G heißt *skalenadäquat von k-ter Ordnung* (kurz: *k-skalenadäquat*), wenn für alle zulässigen Datenmengen $\mathcal{D} \subset \sum_{i\in I}\mathfrak{X}_i$ und alle zulässigen Tupel $T = \left(T_i\right)_{i\in I} \in \prod_{i\in I}\mathbf{T}_i$, die von k-ter Ordnung sind und der Gleichung $A\left(T\right) = Z\left(\mathcal{D}\right)$ genügen, die Inklusion $T = \left(T_i\right)_{i\in I} \in \mathbf{I}_k\left(G_{\mathcal{D}}\right)$ erfüllt ist. Das Hauptproblem für skalenadäquate Zielfunktionen ist dann ebenso kompliziert wie kurz formulierbar. Es lautet einfach: *Bestimme für alle natürlichen Zahlen* $1 \le k \le 3$ *alle k-skalenadäquaten Zielfunktionen* $G : \sum_{i\in I}\mathfrak{X}_i \mapsto \mathbb{R}$!

Eine Beschreibung solcher skalenadäquaten Zielfunktionen zu dem für diesen Abschnitt relevanten Teil ist ebenfalls von HERDEN und PALLACK (2005, S. 24 ff) bereits geliefert worden. Diese Lösung ist als Adäquatheitssatz bekannt.

Angewendet auf die Situation der Leistungsmessung sagt der Adäquatheitssatz im Wesentlichen aus, dass bei Vorlage von endlich vielen Zufallsvariablen, *die nicht notwendig Intervallskalenniveau haben* und einem fest vorgegebenen möglichen Kompromiss X die Funktion G entweder auf der Menge aller Tupel $\left(T\left(X\right), T\left(X_{i_1}\right), ..., T\left(X_{i_n}\right)\right)$, die durch zulässige und das Skalenniveau repräsentierende Transformationen, die nicht notwendig die Form $T\left(x\right) = mx + b$ haben müssen, aus $\left(X, X_{i_1}, ..., X_{i_n}\right)$ hervorgehen, konstant ist, oder aber aus der Menge der Zufallsvariablen $\left\{X, X_{i_1}, ..., X_{i_n}\right\}_{i\in I}$ eine feste Zufallsvariable X_{i_k} und ein festes Kriterium $\omega \in \Omega_{i_k}$ herauswählt, so dass G durch den Wert $X_{i_k}\left(\omega\right)$ bestimmt ist.

Beim Skalenniveau, das nicht durch eine Intervallskala gemessen werden kann, ist das Ergebnis des dem ersten Modell angepassten Adäquatheitssatzes dahin gehend zu interpretieren, dass keine sinnvollen Zielfunktionen existieren. Intervallskalenniveau ist hier allerdings direkt auszuschließen, da es keine Möglichkeit gibt, eine Aussage darüber zu treffen, wie viel wichtiger ein Kriterium als ein zu Vergleichendes ist.

Nichtsdestotrotz wollen wir kurz darauf eingehen, ob es bei Vorlage einer wahren Rangordnung $\omega_1 \precsim \omega_2 \precsim ... \precsim \omega_n$ der Kriterien möglich ist, diese wahre Rangordnung durch Paarvergleiche zu gewinnen. Das bedeutet, dass die einzelnen Kriterien paarweise miteinander verglichen werden. Dazu muss ein Ähnlichkeitsmaß $s : \mathfrak{X} \times \mathfrak{X} \mapsto \left[-1,1\right]$ gefunden werden, welches die wahren Ähnlichkeiten der Kriterien wiedergibt. Damit ein solches Ähnlichkeitsmaß s durch Paarvergleiche überhaupt in der Lage ist, die wahre Rangordnung zu finden, muss s mindestens die beiden folgenden Bedingungen erfüllen.

B1: Wenn $\omega_i \sim \omega_j$, so folgt $s\left(\omega_i, \omega_k\right) = s\left(\omega_j, \omega_k\right)$.

B2: Wenn $\omega_i \prec \omega_j \prec \omega_k$, so folgt $s\left(\omega_i, \omega_j\right) < s\left(\omega_i, \omega_k\right)$ und $s\left(\omega_j, \omega_k\right) < s\left(\omega_i, \omega_k\right)$.

Diese Bedingungen meinen inhaltlich, dass näher beieinanderliegende Paare eine größere Ähnlichkeit aufweisen, als weiter auseinanderliegende Paare.

Unter der Voraussetzung, dass nicht notwendigerweise Intervallskalenniveau vorliegt, lässt sich beweisen, dass derartige Ähnlichkeitsmaße nicht existieren. Die Gewinnung der Rangordnung durch Paarvergleiche ist vollkommen ausgeschlossen (vgl. HERDEN; PALLACK 2005, S. 33).

Die beiden Aussagen des ersten Modells lassen folglich keinerlei Möglichkeiten zu, durch Optimierung von Zielfunktionen eine geeignete Rangfolge der vorliegenden Kriterien zu finden oder die mögliche *wahre* Reihenfolge der Kriterien durch Paarvergleiche der Experten zu rekonstruieren. Daher versuchen wir nun das Modell zu vereinfachen, indem wir die Suche nach einer geeigneten Zielfunktion, deren optimale Lösungen entsprechende Rangfolgen festlegt, aufgeben, um lediglich festzustellen, ob stabile Rangordnungen existieren. Zudem betrachten wir nun, wie in der Praxis üblich, nur endlich viele Kriterien.

Für das zweite Modell gehen wir von einer abzählbar unendlichen Familie $\left(\Omega_i, \mathcal{A}_i, P_i\right)_{i \in I}$ von verschiedenen endlichen Wahrscheinlichkeitsräumen aus, wobei Ω_i die in einer bestimmten Situation betrachtete Menge von Kriterien zur Leistungsmessung repräsentiert. Für diese Familie sollen folgende Bedingungen erfüllt sein:

C1: Die einzelnen Ω_i enthalten mindestens drei Elemente.

C2: Die Familie $\left(\Omega_i, \mathcal{A}_i, P_i\right)_{i \in I}$ soll vereinigungsstabil sein. Das bedeutet, dass zu je zwei Wahrscheinlichkeitsräumen $\left(\Omega_{i_1}, \mathcal{A}_{i_1}, P_{i_1}\right)$ und $\left(\Omega_{i_2}, \mathcal{A}_{i_2}, P_{i_2}\right)$ ein Wahrscheinlichkeitsraum $\left(\Omega_{i_3}, \mathcal{A}_{i_3}, P_{i_3}\right)$ existiert, so dass $\left(\Omega_{i_3}, \mathcal{A}_{i_3}, P_{i_3}\right)_{i \in I} = \left(\Omega_{i_1} \cup \Omega_{i_2}, \mathcal{A}_{i_3}, P_{i_3}\right)$ gilt.

C3: Für jedes Ω_i und jede mindestens dreielementige Teilmenge T von Ω_i, existiert ein $j \in I$ mit $\Omega_j = T$.

Wir betrachten auf jedem dieser endlichen Wahrscheinlichkeitsräume die Menge $\mathfrak{X}_i$ aller reellen Zufallsvariablen. Diese Zufallsvariablen repräsentieren erneut die Experten und werden in gleicher Weise interpretiert wie in dem ersten Modell. Unterstellt man, dass die Zufallsvariablen kein Intervallskalenniveau haben, so folgt daraus wegen der Endlichkeit der Wahrscheinlichkeitsräume, dass nur ordinale Signifikanz der Zufallsvariablen vorliegt (vgl. HERDEN; PALLACK 2005).

Im Unterschied zum ersten Modell sind wir nicht mehr an Gütekriterien interessiert, sondern wir betrachten Rangfunktionen $r_i : \Omega_i \mapsto \{1, \ldots, |\Omega_i|\}$. Die Rangfunktionen sollen die Meinungen der Experten zu einem Rang zusammenfassen. Eine Funktion von Ω_i heißt Rangfunktion, wenn sie bijektiv bzw. lückenlos ist und gleichzeitig zwei Kriterien erfüllt:

K1: Wenn für alle $X, X' \in \mathfrak{X}_i$ und für alle $\omega, \omega' \in \Omega_i$ die Äquivalenz $X(\omega) \leq X(\omega') \Leftrightarrow X'(\omega) \leq X'(\omega')$ gilt, soll daraus $r(\omega) \leq r(\omega')$ folgen.

Inhaltlich bedeutet dies, dass unter der Voraussetzung, dass sich alle Experten über die Rangfolge völlig einig sind, die Rangfolge der Experten natürlich auch die richtige Rangfolge sein sollte.

Das zweite Kriterium, die Stabilität einer Familie $(r_i)_{i \in I}$ von Rangfunktionen betreffend, bedarf zunächst einer Definition des Begriffs der *Erweiterung* einer Zufallsvariablen X, die auf der Menge Ω_i definiert ist, zu einer Zufallsvariablen X', die auf einer Ω_i umfassenden Menge Ω_j definiert ist.

Definition:
Seien $X \in \mathfrak{X}_i$ eine Zufallsvariable, die auf dem Wahrscheinlichkeitsraum $(\Omega_i, \mathcal{A}_i, P_i)_{i \in I}$ definiert ist und $Y \in \mathfrak{X}_j$ eine Zufallsvariable, die auf dem Wahrscheinlichkeitsraum $(\Omega_j, \mathcal{A}_j, P_j)_{j \in I}$ definiert ist, wobei $\Omega_i \subset \Omega_j$ gelte. Dann heißt Y *Erweiterung* von X genau dann, wenn für alle $\omega, \omega' \in \Omega_i$ die Äquivalenz $X(\omega) \leq X(\omega') \Leftrightarrow Y(\omega) \leq Y(\omega')$ gilt. Die Erweiterungen sind nur bis auf Ordnungsautomorphismen bestimmt. Dies bedeutet, dass eingeschränkt auf Ω_i Y mit X bis auf Ordnungsautomorphismen übereinstimmt.

Nun kann das zweite Kriterium formuliert werden.

K2: Eine Familie $\left(r_i\right)_{i\in I}$ von Rangfunktionen heißt stabil, wenn sie das Kriterium **K1** erfüllt und darüber hinaus für alle $i \in I$ und alle $j \in I$, wobei weiterhin $\Omega_i \subset \Omega_j$ gelte, und für alle $\omega, \omega' \in \Omega_i$ die Äquivalenz: $r_i\left(\omega\right) \le r_i\left(\omega'\right) \Leftrightarrow r_j\left(\omega\right) \le r_j\left(\omega'\right)$ gilt.

Dies bedeutet inhaltlich, dass sich die Rangfolge von Kriterien bei Erweiterung oder Einschränkung der Kriterien nicht ändern soll. Die Experten sollen in ihrer Meinung folglich zumindest so stabil sein, dass sie die Reihenfolge, die sie bei den noch übrig gebliebenen Kriterien schon vorher gewählt hatten, auch dann noch beibehalten, wenn einzelne Kriterien gestrichen oder hinzugefügt werden.

Leider führt die Suche nach stabilen Rangordnungen unter den gegebenen Voraussetzungen zu einem Diktator-Theorem vom Arrow-Typ: Eine Familie $\left(r_i\right)_{i\in I}$ von Rangordnungen ist genau dann stabil, wenn sie im Wesentlichen mit einer Zufallsvariablen X_i und deren Erweiterungen (bzw. Einschränkung) übereinstimmt. Das bedeutet, dass zu dieser Familie $\left(r_i\right)_{i\in I}$ von Rangfunktionen eine Familie von Zufallsvariablen existiert, wobei die einzelnen Zufallsvariablen eingeschränkt auf eine feste Menge bis auf ordinale Signifikanz eindeutig bestimmt sind und darüber hinaus zu jeder Menge Ω_i eine Zufallsvariable X_i dieser Familie existiert, so dass für alle $\omega, \omega' \in \Omega_i$ folgende Äquivalenz gilt: $r_i\left(\omega\right) \le r_i\left(\omega'\right) \Leftrightarrow X_i\left(\omega\right) \le X_i\left(\omega'\right)$ (vgl. Herden 1999, S. 129 f).

Dies impliziert sofort, dass Rangfunktionen, die man als Konsens bezeichnen könnte, ebenfalls nicht existieren, weil ein *Diktator* eine solche konsistente Rangfunktion durch eine einzige Funktion und ihre Erweiterung bestimmt. Folgerichtig bedeutet dies, dass jeder Kompromiss zwischen verschiedenen Gruppen von Individuen mit unterschiedlichen Interessen bzw. Schwerpunktsetzungen nicht stabil sein kann. Sobald zwei verschiedene Experten zwei verschiedene Meinungen haben, kann es keine stabile Rangfunktion geben, welche die obigen Kriterien erfüllt. Ein Konsens ist dementsprechend auch im zweiten Modell nicht möglich.

Eine detailliertere Ergebnisdarstellung ist bei Herden (1999) nachzulesen. Dazu vergleiche man insbesondere Satz 1, Satz 2 und die zugehörigen Kommentare. Es ist allerdings zu beachten, dass Herden, im Gegensatz zu dem hier vorgestellten Modell, die Menge Ω_i als Menge der Experten und die Menge der reellen Zufallsvariablen $\mathfrak{X}_i$ als Menge der Kriterien interpretiert. Mathematisch macht dies keinen Unterschied, so dass wir das Modell benutzen können.

Während das erste Modell für die praktische Vorgehensweise vollständig ausscheidet, legt das zweite Modell zumindest die Möglichkeit nahe, dass man durch eine Art vollständigen Kompromiss den geeigneten Diktator finden kann, welcher die Rangfolge der Kriterien zur Leistungsmessung festlegt. Zur mathematischen Approximation dieses Problems, der Suche nach dem geeigneten Diktator, kann man die, sich allerdings noch im Anfangsstadium befindliche, Theorie aus der von HERDEN betreuten Dissertation von BECKE aus dem Jahre 1996 heranziehen, in der der Begriff des Konsenskriteriums definiert ist.

Umgangssprachlich versteht man unter einem Konsens die *„Übereinstimmung der Meinungen"* (Duden o.J.), der an einer Diskussion beteiligten Akteure. Bei BECKE (1996) wird diese gemeinsame Meinung durch die ermittelte gemeinsame Rangfolge der Kriterien repräsentiert, während die ursprünglichen Einzelmeinungen der beteiligten Akteure durch die von den Experten vergebenen Rangordnungen wiedergegeben werden. Häufig weichen die Einzelmeinungen allerdings stark von der gemeinsamen Meinung ab, so dass der Grad des Konsens innerhalb der betrachteten Expertenmenge nicht besonders hoch ist, wie nachfolgendes Beispiel illustriert:
Entscheidet sich bei der Diskussion um die mathematische Leistungsmessung die eine Hälfte der Experten dafür, dass „Lesekompetenz" ein sehr wichtiges mathematisches Kriterium repräsentiert, während die andere Hälfte dieses Kriterium als äußerst unwichtig ansieht, ist es sicherlich problematisch, eine durchschnittliche Beurteilung des Kriteriums als sinnvollen Konsens zu bezeichnen. Plausibler erscheint es in diesem Fall festzuhalten, dass keinerlei Konsens vorliegt, da die möglichst große Übereinstimmung der Experten als Kriterium für einen Konsens offensichtlich nicht vorhanden ist.

Dementsprechend erfüllt eine Menge von Experten ein Konsenskriterium nach BECKE (1996, S. 10 ff) umgangssprachlich genau dann, wenn die Meinungen der einzelnen Experten möglichst wenig differieren. Ein Konsenskriterium dient also vor allem der Beurteilung einer möglicherweise konsensfähigen Rangordnung. Bei zu starkem Auseinanderklaffen der Expertenmeinungen ist es nämlich durchaus möglich, dass überhaupt kein Konsens existiert.

Im Sinne einer mathematischen Präzision betrachten wir erneut ein einfaches wahrscheinlichkeitstheoretisches Modell mit einem einzigen endlichen Wahrscheinlichkeitsraum $\left(\Omega, \mathcal{A}, P\right)$ und einer Menge $\mathfrak{X}$ von reellen Zufallsvariablen, welche die Experten repräsentiert. Ω enthält n Kriterien zur Leistungsmessung. Wir nehmen nun an, dass ein Kompromiss gefunden wurde, also irgendeine vorgegebene Rangfunktion $r : \Omega \mapsto \left\{1, \ldots, n\right\}$ existiert. Wir wollen nun untersuchen, ob diese Rangfunktion r einen *sinnvollen* Konsens repräsentiert, bei der also die subjektiven Einzelmeinungen der Experten nicht zu weit differieren. Dazu nehmen wir an, dass die vorgegebene Rangfunktion r die *wahre* Rangfunktion sei, und vergleichen diese mit den einzelnen Rangfunktionen der Experten.

Dann ist jedem Kriterium $\omega_i \in \Omega$ ein *wahrer* Rang $r\left(\omega_i\right) \in \left\{1,\dots,n\right\}$ und eine Fehlervariable $F_i : \mathfrak{X} \mapsto \left\{1 - r\left(\omega_i\right),\dots,n - r\left(\omega_i\right)\right\}$ zugeordnet (für $i \neq j$ gelte $r\left(\omega_i\right) \neq r\left(\omega_j\right)$ $\left(1 \leq i, j \leq n\right)$). Weicht nämlich die subjektive Meinung der Experten von dem *wahren* Rang ab, unterlaufen den Experten *Fehler*, weil der Rang entweder als zu niedrig (positiver Fehler) oder als zu hoch (negativer Fehler) eingeschätzt wird. Der *Fehler* $X_i\left(\omega\right) - r\left(\omega\right)$ ist also die Abweichung des Rangs $X_i\left(\omega\right)$, den der Experte $X_i \in \mathfrak{X}$ einem Kriterium $\omega \in \Omega$ zugewiesen hat, von dem *wahren* Rang $r\left(\omega\right)$ des Kriteriums. Geht man nun aber davon aus, dass es sich um Experten handelt, kann nur dann von einem vernünftigen Konsens gesprochen werden, wenn den Experten im Gruppenmittel absolut größere Fehler mit nicht größerer Wahrscheinlichkeit unterlaufen als absolut kleinere. Zur mathematischen Präzisierung dieser Forderung sei für alle $i \in \left\{1,\dots,n\right\}$ die Menge $\left\{F_i = k\right\}$ durch $\mathcal{A}_i^k$ abgekürzt. Das oben für einen sinnvollen Konsens geforderte Verhalten der Experten wird dann bei Becke (1996, S. 30 f) folgendermaßen präzisiert:

$$\forall i \in \left\{1,\dots,n\right\} \forall j \in \left\{1,\dots,n\right\} \forall k \in \operatorname{Im}\left(F_i\right) \forall l \in \operatorname{Im}\left(F_j\right) :$$
$$\left(|\, k \,|<|\, l \,| \Rightarrow P\left(\mathcal{A}_i^k\right) \geq P\left(\mathcal{A}_j^l\right)\right).$$

Falls ein solches Konsenskriterium nicht erfüllt ist, ist es möglich Kriterien zusammenzufassen, indem einzelne Kriterien, die untereinander offensichtlich wenig, aber von anderen Kriterien stärker abgrenzbar sind, zu Gruppen zusammengefasst werden. Nach der durchgeführten Zusammenfassung zu Gruppen besteht über die Gruppen eher ein Konsens als über die vorher betrachteten einzelnen Kriterien. Diese Zusammenfassung geschieht so lange, bis eine konsensfähige Rangordnung ermittelt werden kann. Dieser *Erfüllbarkeitsgrad* ist ein Kriterium dafür, inwieweit die Gruppe der Experten konsensfähig gewesen ist (vgl. ebd., S. 14).

Zum besseren Verständnis betrachte man folgendes Beispiel von Becke (1996, S. 36 ff), welches der Thematik der Leistungsmessung angepasst worden ist:

Man nehme an, dass sich acht Didaktiker auf sieben fundamentale Kriterien zur Leistungsmessung geeinigt haben. Zur Ermittlung der Rangfolge dieser Kriterien habe jeder Experte die Kriterien entsprechend der subjektiv eingeschätzten Bedeutung nach geordnet, wobei das jeweils wichtigste Kriterium an erster Stelle steht. Diese Kriterien lauten:

ω_1 : Logisches Denkvermögen

ω_2 : Mengen- und Zahlverständnis

ω_3 : Räumliche Vorstellungskraft

ω_4 : Mathematische Modellierung

ω_5 : Abstraktionsvermögen

ω_6 : Motivation

ω_7 : grundlegende Rechenfertigkeiten

Die Indizes der Kriterien geben den *wahren* Rang des zugehörigen Kriteriums an. Die acht Experten $X_1, ..., X_8$ vergaben folgende Ränge:

1. Experte (X_1)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	1	3	7	2	4	5	6

2. Experte (X_2)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	2	4	1	3	5	7	6

3. Experte (X_3)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	4	1	3	2	5	6	7

4. Experte (X_4)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	1	3	2	4	5	6	7

5. Experte (X_5)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	2	1	3	4	5	7	6

6. Experte (X_6)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	1	2	3	4	5	6	7

7. Experte (X_7)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	1	3	2	5	4	6	7

8. Experte (X_8)

Kriterium	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
Rang	1	4	2	5	3	7	6

Die aus diesen Daten errechneten arithmetischen Mittelwerte $E_p(\omega_i)$ $(1 \leq i \leq 7)$ der Kriterien $\omega_1, \ldots, \omega_7$ ergeben dann in aufsteigender Größe die Rangfolge der Kriterien. Wir nehmen an, dass es sich dabei um die *wahre* Rangfolge handelt, und untersuchen nun, ob kleinere Fehler mit größerer Wahrscheinlichkeit begangen werden als große Fehler.

Man betrachte nun die Tabellen der absoluten Fehlerhäufigkeiten:

Logisches Denkvermögen

Fehler k	0	1	2	3	4	5	6
Häufigkeit	5	2		1			

Mengen- und Zahlverständnis

Fehler k	-1	0	1	2	3	4	5
Häufigkeit	2	1	3	2			

Räumliche Vorstellungskraft

Fehler k	-2	-1	0	1	2	3	4
Häufigkeit	1	3	3				1

Mathematische Modellierung

Fehler k	-3	-2	-1	0	1	2	3
Häufigkeit		2	1	3	2		

Abstraktionsvermögen

Fehler k	-4	-3	-2	-1	0	1	2
Häufigkeit				2	5	1	

Motivation

Fehler k	-5	-4	-3	-2	-1	0	1
Häufigkeit					1	4	3

grundlegende Rechenfertigkeiten

Fehler k	-6	-5	-4	-3	-2	-1	0
Häufigkeit						4	4

Bereits ein flüchtiger Blick auf die abgebildeten Tabellen verdeutlicht, dass das vorgeschlagene Konsenskriterium nicht in *ausreichendem* Maße erfüllt ist, so dass auf Grundlage der Umfrage hier kein vernünftiger Konsens zwischen den Experten festzustellen ist.

In der Dissertation von Becke (1996) ist das obige vorgestellte Konsenskriterium nur eines einer Vielzahl möglicher Konsenskriterien. Tatsächlich hat Becke (1996) in seiner Dissertation einen sehr abstrakten, sehr allgemeinen und sehr komplizierten wahrscheinlichkeitstheoretisch orientierten Begriff eines allgemeinen Konsenskriteriums entwickelt. Die daraus resultierende sehr komplizierte Theorie hat allerdings den Vorteil, dass tatsächlich ein (noch nicht programmierter) Algorithmus entwickelt worden ist, der es ermöglicht bei Vorlage eines Konsenskriteriums die dazu gehörige Rangfunktion, wenn sie denn existiert, zu finden, für die das entsprechende Konsenskriterium erfüllt ist. Dies kann aber dazu führen, dass keine oder sogar mehrere Rangfunktionen existieren. Außerdem ist nicht eindeutig bestimmt, welches Konsenskriterium verwendet werden soll, da die Theorie darüber bisher keine Aussagen zu treffen in der Lage ist. Hinzu kommt, dass diese Kriterien in der Regel in der Praxis nicht erfüllt sind, was zu Zusammenfassungen der subjektiven Rangordnungen der Experten führen muss, so dass wenigstens ein optimaler Grad an Erfüllbarkeit garantiert ist (vgl. die entsprechende von Becke (1996) entwickelte Theorie des Erfüllbarkeitsgrads).

Im Rahmen dieser Arbeit ist dieser Ansatz, so vielversprechend er sein mag, somit auch nicht a priori tragbar, und soll daher nicht weiter verfolgt werden.

7.3 Leistungsmessung bei vorliegenden Kriterien

Um nun zu modellieren, ob Leistungsmessung im konkreten Fall bei vorliegenden Kriterien möglich ist, orientiert sich das Vorgehen in diesem Abschnitt im Wesentlichen an zwei Modellen. Das zweigeteilte erste Modell ermittelt die Leistung bei Vorlage von Testergebnissen durch die *Optimierung einer Zielfunktion* und wird auf das erste vorgestellte Modell von Herden und Pallack von 2005 zurückgeführt. Abschließend wird die populäre probabilistische Testtheorie kurz vorgestellt und erklärt, warum sich dieses Vorgehen im Rahmen dieser Arbeit nicht eignet.

Entgegen der Ausführungen in Abschnitt 7.2 nehmen wir für das erste Modell an, dass tatsächlich ein Konsens über die Kriterien zur Leistungsmessung existiert und sich diese Kriterien zusätzlich durch einzelne Aufgaben abprüfen lassen

Dazu betrachten wir zunächst erneut ein wahrscheinlichkeitstheoretisches Modell und den Wahrscheinlichkeitsraum $\left(\Omega,\mathcal{A},P\right)$. Die große Urne Ω beinhaltet nun beliebig viele Probanden, deren Leistung gemessen werden soll, während die auf dem Wahrscheinlichkeitsraum definierte endliche Menge $\mathfrak{X}$ von reellen Zufallsvariablen die n verschiedenen Aufgaben repräsentiert, welche geeignet seien, Leistung zu messen und die Leistungskriterien erfüllen. $X_i(\omega)$ ist dann derjenige *Leistungs*grad, den der Proband $\omega \in \Omega$ bei der Aufgabe i $(1 \leq i \leq n)$ zugewiesen bekommt.

Unter den obigen Voraussetzungen entspricht die Leistungsmessung einer Zielfunktion $G:\left\{\left(X_1(\omega),\ldots,X_n(\omega)\right) \middle| \omega \in \Omega\right\} \mapsto \mathbb{R}$, die jedem Tupel $\left(X_1(\omega),\ldots,X_n(\omega)\right)$ eine bestimmte Leistung zuordnet, indem sie die bei den verschiedenen Aufgaben $X_1,\ldots,X_n$ erreichten Leistungen zusammenfasst. In der Schule werden dazu in der Regel die erzielten Punkte addiert und zu einer Zensur gruppiert. Es stellt sich nun die Frage, ob auf Grundlage dessen eine Leistungsmessung im konkreten Fall tatsächlich möglich ist.

Dazu unterstellen wir, dass die Bewertung nicht sicher auf Intervallskalenniveau gemessen werden kann. Es ist also nicht möglich genau anzugeben, wie viel schwieriger/wichtiger die Aufgabe X_i im Vergleich zur Aufgabe X_j ist, da ein Gutachter zwar wissen kann, ob eine Aufgabe schwieriger oder wichtiger für das mathematische Verständnis der Probanden ist als eine andere Aufgabe, jedoch nicht genau angeben kann, wie viel schwieriger/wichtiger.

Somit hat jedes solche Tupel $\left(X_1(\omega),\ldots,X_n(\omega)\right)$ lediglich ordinale Signifikanz, so dass das erste wahrscheinlichkeitstheoretische Modell aus Abschnitt 7.2 erneut greift. Das bedeutet, wenn man Konstanz der Funktion G ausschließt, dass bei einem einzelnen Probanden ω eine Aufgabe $X_i \in \mathfrak{X}$ ausgewählt wird und die bei dieser Aufgabe erbrachte Leistung $X_i(\omega)$ die ganze Zensur bedingen würde. Das ist natürlich vollkommen unsinnig und spiegelt die vom Probanden erbrachte Leistung in keiner Weise wieder! Wenn also Unsicherheit über die genaue Schwierigkeit oder die genaue Wichtigkeit einer Aufgabe besteht, so ist eine Leistungsbeurteilung unmöglich.

Nach diesem negativen Ergebnis beschäftigen wir uns kurz mit einer anderen, in der Schule eher unüblichen, Form der Leistungsmessung. Dazu betrachten wir abermals den Wahrscheinlichkeitsraum $(\Omega,\mathcal{A},P)$, wobei die Urne Ω sowie die endliche Menge $\mathfrak{X}$ von reellen Zufallsvariablen so wie oben interpretiert werden. Wir versuchen nun eine reelle Zufallsvariable X zu finden, die ein vorgegebenes Tupel $\left(X_1(\omega),\ldots,X_n(\omega)\right)$ optimal repräsentiert, also die erbrachte Leistung optimal zusammenfasst. Das läuft darauf hinaus, dass $G\left(X(\omega)\right)$ minimal ist unter allen möglichen Bewertungen. G ist dabei diejenige Funktion, die angibt, wie gut $X(\omega)$ die Gesamtbewertung zusammenfasst. Die bestmögliche Zusammenfassung der Gesamtbewertung erhält man dann, wenn G minimiert wird (vgl. das unter 7.2 beschriebene

Modell). Wenn wir voraussetzen, dass die Zufallsvariablen lediglich ordinale Signifikanz haben, greift erneut das für das erste Modell gewonnene Resultat und wir haben keine Möglichkeit der Leistungsmessung.

Zusammengefasst bedeutet dies, dass diese Funktionen nicht geeignet sind, Leistung zu messen, wenn die betrachteten Zufallsvariablen lediglich feststellen, ob eine Aufgabe schwieriger ist als eine andere, nicht aber wie viel genau. Jede Bewertung, die ein Gutachter durchführt, muss folglich davon ausgehen, dass die Aufgaben mindestens auf Intervallskalenniveau gemessen werden. Wie oben bereits angedeutet, ist eine solche Bedingung aus verschiedenen Gründen allerdings äußerst fragwürdig. Häufig wird der Schwierigkeitsgrad einer Mathematikaufgabe anhand der relativen Lösungshäufigkeiten ermittelt. Die Verwendung der relativen Lösungshäufigkeiten als Indiz für den Schwierigkeitsgrad einer Mathematikaufgabe wird meist durch das schwache (oder starke) Gesetz der großen Zahlen gerechtfertigt. Diese Rechtfertigung setzt voraus, dass die Erwartungswerte der die Probanden repräsentierenden Zufallsvariablen gleich sind. Da aber davon auszugehen ist, dass der subjektive Schwierigkeitsgrad einer Aufgabe in der Regel, je nach Vorkenntnis, Begabung, Ausbauer usw., von Proband zu Proband variieren wird, ist die Annahme der Gleichheit entsprechender Erwartungswerte zumindest sehr fragwürdig. Des Weiteren setzt dieses Vorgehen, um als objektives Kriterium zu dienen, riesige Probandenmengen (in der schulischen Realität völlig unmöglich) voraus. Es ist nicht einmal auszuschließen, dass Aufgaben schwieriger sein können bzw. schlechter ausfallen, obwohl sie für den Aufbau des Stoffes und die Entwicklung mathematischer Kompetenzen weniger wichtig sind. Dann würde man bei der Leistung nicht die Wichtigkeit für die Mathematik messen. Wir haben also in diesem Fall keine Chance Leistung zu messen, da sich die Bewertungen in der Regel nur auf ordinalem Niveau bewegen.

Einige weitere Varianten zur Messung von Mathematikleistungen von Schülerinnen und Schülern bei vorliegenden Kriterien sind unter dem Oberbegriff **Probabilistische Testtheorie** zu finden. Unter dem Sammelbegriff sind verschiedene Testmodelle zusammengefasst, die Annahmen über den Zusammenhang zwischen den Antworten auf Items (hier: Mathematikaufgaben) und den dazugehörigen latenten Variablen (hier: Mathematikleistung) aufstellen. Das Besondere dabei ist, dass die Beziehung zwischen Antwort und latenten Variablen nicht deterministisch, sondern messtheoretischer Natur ist, da *„die Wahrscheinlichkeit einer bestimmten Antwort auf eine Testaufgabe [ ] als mathematische Funktion von Personenmerkmal(en) und Itemmerkmal(en) definiert"* (Walter 2013) wird. Um derartige Modelle zur Leistungsmessung nutzen zu können, müssen in der Regel allerdings zwei wesentliche Voraussetzungen erfüllt sein, ohne deren Einhaltung ein nicht zu vertretender Informationsverlust möglich ist.

(i) Eindimensionalität: Nur wenn genau eine latente Variable (hier: Mathematikleistung) das Antwortverhalten bei einem Item (hier: Testaufgabe) bestimmt, können probabilistische Testmodelle verwendet werden. Keine weitere latente Variable darf einen systematischen Einfluss ausüben (vgl. WEBER 2005).
In der Praxis der mathematikdidaktischen Forschung werden die latenten Variablen, mit denen Mathematikleistung beschrieben wird, ähnlich wie die latenten Variablen für Intelligenz, in der Regel durch eine Faktorenanalyse ermittelt. Dieses Vorgehen wird beispielsweise bei der PISA-Studie verwendet. Dazu geht man davon aus, dass die zur Verfügung stehenden Testaufgaben geeignet sind, Mathematikleistung zu messen. Mit Hilfe einer explorativen Faktorenanalyse der Testaufgaben wird dann untersucht, wie viele latente Variablen benötigt werden, um die Mathematikleistung von Schülerinnen und Schülern zu bestimmen. Bei der PISA-Studie ist dabei genau eine latente Variable extrahiert worden, mit der die Lösungen in Mathematikaufgaben erklärt werden können, so dass probabilistische Modelle Anwendung finden können (vgl. Deutsches PISA-Konsortium 2005).

(ii) (lokale) stochastische Unabhängigkeit: Die zweite zentrale Voraussetzung beschreibt im Wesentlichen, dass die Korrelation zwischen den Testaufgaben einzig durch die latente Variable bestimmt ist: Die Lösungswahrscheinlichkeit eines Schülers bei einer Aufgabe hängt mit den Fähigkeiten des Schülers zusammen und nicht mit der richtigen oder falschen Beantwortung einer vorher bearbeiteten Aufgabe. Die Beobachtungen müssen also unabhängig voneinander gemacht werden und dürfen sich gegenseitig nicht beeinflussen. Testaufgaben, die aufeinander aufbauen, verletzten diese Annahme beispielsweise (vgl. WALTER 2013). Zudem ist diese Unabhängigkeitsforderung bereits bei elementaren mathematischen Fertigkeiten nicht erfüllt. LIND (2002, S. 7) fällt

> *„zum Beispiel bei* 1×1 *-Aufgaben auf, dass es Schüler gibt, die unsicher bei der Multiplikation mit* 0 *sind. Ein solcher Schüler kann bei den restlichen Aufgaben durchaus 'besser' als ein Mitschüler* m *sein, obwohl er* m *bei der Multiplikation mit* 0 *unterlegen ist".*

Nur wenn diese Voraussetzungen erfüllt sind, kann man versuchen, mit Hilfe probabilistischer Testmodelle auf Grundlage der gewonnenen Testergebnisse auf die zugrunde liegende Mathematikleistung zurückzuschließen.

Das bekannteste und mathematisch-statistisch fundierteste Modell ist das auf den dänischen Statistiker GEORG RASCH (*1901; †1980) zurückgehende Rasch-Modell von 1960, welches häufig angewendet wird, wenn es um die Untersuchung eindimensionaler Persönlichkeitsmerkmale geht (vgl. LIND 1994, S. 280). Das Modell definiert die Wahrscheinlichkeit einer Antwort eines Probanden v auf ein dichotomes (als falsch oder richtig gelöst codiertes) Item als logistische Funktion aus Personenfähigkeit ξ_v und Schwierigkeit der Aufgabe σ_i (vgl. WALTER 2013). Die Modellgleichung lautet $P\left(A_{vi}=1\right)=\frac{e^{\xi_v-\sigma_i}}{1+e^{\xi_v-\sigma_i}}$ (vgl. PPT (o.V.) o.J., S. 4). Der Gleichung kann beispielsweise entnommen werden, dass ein Anstieg der Aufgabenschwierigkeit mit einer Abnahme der Wahrscheinlichkeit, die Aufgabe korrekt lösen zu können, verbunden ist. Wenn das Rasch-Model angewendet werden soll, müssen die Personenparameter mindestens einer Intervallskala entsprechen.

Die gesamte Information über die Mathematikleistung jedes Probanden wird schließlich als Summe (Summenscore) aus allen Einzelaufgaben bestimmt. Während dieses Prinzip im Modell der klassischen Testtheorie bereits enthalten ist, geht es im Rasch-Modell speziell als Annahme mit ein. Die Anzahl gelöster Aufgaben ist eine erschöpfende Statistik für den Fähigkeitsparameter der Schülerin bzw. des Schülers und gibt letztlich die Mathematikleistung an (vgl. ebd., S. 4). Diese Eigenschaft wird als Suffizienz bezeichnet (vgl. WALTER 2013).

Eine besondere Leistung dieses Modells ist laut LIND (1994, S. 279 ff) zudem die Möglichkeit, die Schwierigkeitsparameter der Testaufgaben unabhängig von den Probandenfähigkeiten schätzen zu können (Separabilität). Zudem geschieht die Parameterschätzung der Fähigkeiten und Schwierigkeiten unabhängig von der Aufgabenstichprobe und der Personenstichprobe, wobei die Genauigkeit der Schätzung natürlich mit wachsender Stichprobengröße zunimmt (spezifische Objektivität). Die Parameter der logistischen Funktion des Rasch-Modells werden durch eine bedingte Maximum-Likelihood-Schätzung ermittelt (vgl. PPT (o.V.) o.J., S. 5).

Werden die Modellannahmen des Rasch-Modells allerdings verletzt, ist die Verwendung nicht angezeigt. Im Rahmen dieser Arbeit muss davon ausgegangen werden, dass die Voraussetzungen nicht erfüllt werden können. Zunächst ist es dem Verfasser nicht möglich, die Eindimensionalität von Mathematikleistung zu garantieren. Das würde bedeuten, dass nur ein einziges Kriterium existiert, was die bisherige Modellbildung in diesem Kapitel extrem einschränken würde. Zudem ist bei nicht vorliegender Eindimensionalität nicht auszuschließen, dass die n latenten Variablen zwar intervallskaliert sind, aber die Intervallskalen, auf denen diese n latenten Variablen gemessen werden, unterschiedlicher Natur sind. Dann müssen Transformationen der Daten zugelassen sein, die den einzelnen Skalen angepasst sind. Wenn dies aber der Fall ist, gibt es nach Satz 2.4 von HERDEN (1990, S. 85) keine vernünftigen Gütekriterien für derartig skalierte Daten. Dies verdeutlicht die grundsätzliche Problematik bei verletzter Eindimensionalität und u. U. intervallskalierten Daten.

Ein zweites Problem ist sicherlich, dass die gesamte Information über die Leistung der Probanden in den Summenscores festgehalten wird. Auch diese explizite Annahme des Rasch-Modells können wir nicht rechtfertigen, insbesondere dann nicht, wenn die Testergebnisse entsprechend der obigen Ausführungen nicht mindestens auf Intervallskalenniveau gemessen werden. Während das Rasch-Modell Intervallskalenniveau voraussetzt, ist es das Ziel dieser Arbeit, Verfahren zu entwickeln, die lediglich das in die Datenauswertung mit einbringen, was tatsächlich beobachtbar ist. Dabei handelt es sich in der Regel um beobachtbare „je-desto"-Aussagen.

Schließlich sei noch kurz die Frage diskutiert, inwieweit sich nicht intervallskalierte Daten durch geeignete Transformationen in Daten überführen lassen, die intervallskaliert sind. Dazu betrachten wir eine Zufallsvariable $X: \Omega \mapsto \mathbb{R}^{>0}$, die nicht intervallskaliert ist, und einen zulässigen Ordnungsisomorphismus $\varphi : (\mathbb{R}^{>0}, \leq) \mapsto (\mathbb{R}, \leq)$ (z. B. eine Logarithmusfunktion), so dass $\varphi \circ X$ Intervallskalenniveau besitzen soll. Im folgenden Satz wird gezeigt werden, dass es eine solche Transformation im Allgemeinen nicht gibt, wenn wir unterstellen, dass X auf einer nicht notwendig intervallskalierten Skala gemessen wird. Dieses Ergebnis zeigt nach Meinung des Verfassers die grundsätzliche Problematik sogenannter φ- linearer Funktionen.

Zur Formulierung des entsprechenden Satzes seien $(\mathbb{T}, \circ)$ eine Gruppe von Ordnungsautomorphismen auf $(\mathbb{R}, \leq)$ und $T \in \mathbb{T}$ ein fest gewählter Ordnungsautomorphismus. Dann garantiert das Zorn'sche Lemma die Existenz maximaler Untergruppen von $(\mathbb{T}, \circ)$, die T nicht enthalten. Die Vereinigung dieser Untergruppen sei mit $\mathcal{U}(\mathbb{T}, T)$ abgekürzt. Man verifiziert sofort, dass $\mathcal{U}(\mathbb{T}, T)$ sich aus genau denjenigen Ordnungsautomorphismen $T' \in \mathbb{T}$ zusammensetzt, für die T nicht in der von T' erzeugten Untergruppe von $(\mathbb{T}, \circ)$ enthalten ist.

Seien nun $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $\mathfrak{X}$ eine Menge von reellen Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$, die nicht fast überall konstant sind und deren Messniveau durch eine Gruppe $(\mathbb{T}, \circ)$ von Ordnungsautomorphismen auf $(\mathbb{R}, \leq)$ gegeben sei, d. h. es gilt $\mathfrak{X} = \{T(X) \mid T \in \mathbb{T} \text{ und } X \in \mathfrak{X}\}$. Dann bezeichnen wir mit $\mathbb{T}(T)$ die Menge aller Ordnungsautomorphismen $T \in \mathbb{T}$, die die Eigenschaft haben, dass zu allen Ordnungsautomorphismen $T' \in \mathcal{U}(\mathbb{T}, T)$ und allen Zufallsvariablen $X \in \mathfrak{X}$ eine lineare (affine) Abbildung $\varphi : \mathbb{R} \mapsto \mathbb{R}$ existiert, so dass $T'(T(X)) = \varphi(T(X))$ gilt.

Die aufgeworfene Frage nach Ordnungsautomorphismen, die Zufallsvariable in intervallskalierte Zufallsvariable überführen, lässt sich nun wie folgt präzisieren.

Man bestimme notwendige und hinreichende Bedingungen dafür, dass $\mathbb{T}(T)$ nicht leer ist.

Zur Beantwortung der aufgeworfenen Frage ist folgende Notation von Nutzen:

Sei $T \in \mathbb{T}$ beliebig gewählt, dann heißt $\mathcal{U}(\mathbb{T},T)$ zu einer Menge $\Lambda(\mathbb{T},T)$ von linearen (affinen) Abbildungen auf $\mathbb{R}$ äquivalent, wenn eine bijektive Abbildung $\Psi : \mathcal{U}(\mathbb{T},T) \mapsto \Lambda(\mathbb{T},T)$ existiert so dass für alle $T' \in \mathcal{U}(\mathbb{T},T)$ und alle Zufallsvariablen $X \in \mathfrak{X}$ die Gleichung $T'_{|T\circ X(\Omega)} = \Psi(T')_{|T\circ X(\Omega)}$ gilt.

Mit Hilfe dieser Notation lässt sich folgender Satz, der obige Frage beantwortet, gewinnen.

Satz 7.1:

Folgende Aussagen sind äquivalent:

(i) $\mathbb{T}(T) \neq \varnothing$.

(ii) *Es existiert ein Ordnungsautomorphismus* $T \in \mathbb{T}$, so dass $\mathcal{U}(\mathbb{T},T)$ *zu einer Menge* $\Lambda(\mathbb{T},T)$ *von linearen (affinen) Abbildungen auf* $\mathbb{R}$ *äquivalent ist.*

Beweis:

Da eine lineare (affine) Abbildung $\varphi : \mathbb{R} \mapsto \mathbb{R}$ durch die Vorgabe von zwei Punkten eindeutig bestimmt ist, folgt die behauptete Äquivalenz unmittelbar aus der oben aufgeführten Definition von $\mathbb{T}(T)$.

Q.E.D.

Korollar:

Angenommen, $X(\Omega) = \mathbb{R}$ *für alle* $X \in \mathfrak{X}$. *Dann sind folgende Aussagen äquivalent*:

(i) $\mathbb{T}(T) \neq \varnothing$.

(ii) *Es gibt einen Ordnungsautomorphismus* $T \in \mathbb{T}$, *für den* $\mathcal{U}(\mathbb{T},T)$ *nur aus linearen (affinen) Abbildungen besteht.*

7.4 Die Messung der Mathematikleistung in diesem Forschungsprojekt anhand faktorenanalytischer Überlegungen

In Anbetracht der Tatsache, dass es im Rahmen dieser Arbeit nicht möglich ist, auf Modelle der probabilistischen Testtheorie zurückzugreifen, ist es angezeigt bei der praktischen Vorgehensweise zur Bestimmung der Mathematikleistung der Probanden einen Kompromiss zu versuchen. Dies meint, dass sich der Verfasser ein Stück weit an der üblichen Vorgehensweise orientiert, indem zunächst mit Hilfe der Faktorenanalyse latente mathematikleistungsbe-

schreibende Variablen ermittelt werden. Dennoch unterscheidet sich das Vorgehen bereits an dieser Stelle von der üblichen Handhabung drastisch, da keine standardisierte Form der Faktorenanalyse verwendet wird, um kein Datenniveau unterstellen zu müssen. Daher wird auch an dieser Stelle die in dieser Arbeit entwickelte Form der Faktorenanalyse eingesetzt.
Zusätzlich müssen aber auch die Ergebnisse der Modellierung Beachtung finden. Schließlich ist deutlich geworden, dass man letztlich immer einen *Diktator* braucht. In der Schule gibt es in der Person des Lehrers einen solchen *Diktator*. Aus diesem Grund sollen die Lehrerinnen und Lehrer im Anschluss an die Faktorenanalyse die Mathematikleistungen beurteilen. Diese Art der Leistungsmessung ist weder vollkommen objektiv, noch entspricht sie dem, was man gerne unter Leistungsmessung versteht. Diesem (im Rahmen dieser Arbeit nicht behebbaren) Mangel wird daher bereits im Titel der Dissertation Rechnung getragen, indem von in der Schule gemessenen Mathematikleistungen die Rede ist. Daher bezieht sich die Leistungsmessung der Probanden in dieser Arbeit auf den Anfang der Leistungsmessung durch die Schule. Schließlich beobachten und bewerten die Lehrkräfte ihre Schülerinnen und Schüler täglich. Sie stellen Stärken und Schwächen als Erstes fest. Standardisierte Tests zur Überprüfung dieser Diagnosen werden, wenn überhaupt, erst im Anschluss durchgeführt.

Um nun auf Basis der schulischen Realität die Mathematikleistung der Probanden durch die Lehrkräfte zu messen, sollen, wie eingangs angekündigt, zunächst die latenten Faktoren, die Mathematikleistung beschreiben, ermittelt werden. Dazu werden zunächst 21 Lehrerinnen und Lehrer sämtlicher Schulformen in einem ersten Schritt fragebogenbasiert gebeten, besondere Fähigkeiten zu nennen, die Schülerinnen und Schüler der Sekundarstufe I tatsächlich benötigen, um gute Leistungen in Mathematik zu zeigen. Der konkrete Arbeitsauftrag, der im Anschluss an eine ausführliche Erklärung des Forschungsziels angefügt wird, lautet:

> *„Ich bitte Sie im Folgenden anzugeben, über welche (besonderen) Fähigkeiten Ihrer Meinung nach eine Schülerin/ein Schüler verfügen sollte, damit sie/er in Mathematik gute Leistungen zeigt! Antworten Sie bitte stichwortartig."*

Damit die Lehrkräfte nicht auf eine bestimmte Fährte gelockt werden, sind keine weiteren Instruktionen vorzunehmen, so dass der Fantasie bei der Beantwortung keinerlei Grenzen gesetzt werden. Insgesamt zeigen die Mathematiklehrerinnen und -lehrer über alle Schulformen hinweg ein überraschend homogenes Antwortverhalten. Letztlich kristallisieren sich dreizehn verschiedene Fähigkeiten als besonders bedeutsam heraus.

Dabei handelt es sich um folgende Fähigkeiten:

- Logisches Denkvermögen
- Ausdauer und Fleiß
- Sprachvermögen
- Lesekompetenz

- Räumliche Vorstellungskraft
- Fantasie
- Abstraktionsvermögen
- Gedächtnisleistung
- Motivation (Spaß, Neugier, Interesse)
- Flexibilität
- Mengen- und Zahlverständnis
- Grundlegende Rechenfertigkeiten und sichere Vorkenntnisse
- Mathematische Modellbildung

Die durch diesen Schritt erhaltene Liste wird anschließend der identischen Lehrergruppe mit der Bitte vorgelegt, die dreizehn aufgelisteten Fähigkeiten entsprechend der subjektiven Einschätzung auf einer diskreten Skala mit dreizehn Antwortmöglichkeiten zu bewerten. Auch an dieser Stelle sei der konkrete Arbeitsauftrag kurz angegeben:

> *„Bewerten Sie bitte, welche dieser Fähigkeiten eine gute Mathematikschülerin/einen guten Mathematikschüler besonders stark auszeichnen. '1' entspräche einer weniger existenziellen Fähigkeit; '13' einer einen guten Mathematiker besonders stark auszeichnenden Fähigkeit!"*

Auf Grundlage dieser Bewertungen ist es dem Verfasser möglich, für alle Lehrkräfte die jeweilige individuelle Prioritätsordnung der Fähigkeiten zu ermitteln. Basierend auf dem dadurch gewonnenen und in 0-1-Daten transformierten Datensatz können die der Mathematikleistung zugrunde liegenden latenten Faktoren nun durch das im mathematischen Kapitel vorgestellte und an die übliche Faktorenanalyse angelehnte Verfahren ermittelt werden.

Nach Durchführung sämtlicher relevanter Schritte der Hauptkomponentenmethode mit anschließender Varimax-Rotation (die Vorgehensweise wird in Analogie zu den bisher betrachteten Faktorenanalysen dieser Arbeit weiterhin als bekannt vorausgesetzt) lässt sich die in der Schule von Mathematiklehrerinnen und -lehrern beobachtete Mathematikleistung durch zwei latente Faktoren beschreiben:

(i) *verbale Kompetenzen:* Der erste Faktor umfasst die Fähigkeiten „Sprachvermögen", „Ausdauer/Fleiß", „Phantasie", „Lesekompetenz", „Motivation", „Flexibilität" und „Gedächtnisleistung", die der Verfasser als **verbale Kompetenzen** interpretiert. Die Bezeichnung lässt sich vor allem mit der Dominanz der Fähigkeit „Sprachvermögen" begründen.

(ii) *numerische Intelligenz:* Der zweite Faktor enthält dagegen die Fähigkeiten „Mengen-/Zahlverständnis", „grundlegende Rechenfertigkeit/sichere Vorkenntnisse", „räumliche Vorstellung", „mathematische Modellierung", „logisches Denkvermögen" sowie „Abstraktionsvermögen" und wird vom Verfasser als **numerische Intelligenz** bezeichnet. Die Interpretation des zweiten Faktors basiert vor allem darauf, dass es sich bei den zusammengefassten Fähigkeiten um *typische* und *klassische* Mathematikfertigkeiten handelt.

Die exakte Verteilung der Fähigkeiten auf die beiden latenten Faktoren ist auch der Tabelle 7.1 zu entnehmen:

Tabelle 7.1

Die latenten Faktoren der Mathematikleistung

Verbale Kompetenzen	Numerische Intelligenz
Sprachvermögen Ausdauer/Fleiß Phantasie Lesekompetenz Motivation Flexibilität Gedächtnisleistung	Mengen-/Zahlverständnis Rechenfähigkeit Räumliche Vorstellung Mathematische Modellierung Logisches Denkvermögen Abstraktionsvermögen

Auf Basis der Lehrerumfrage lässt sich Mathematikleistung folglich anhand nachfolgender Definition in der Schule messen:

„Schülerinnen und Schüler, die gute Leistungen in Mathematik zeigen, verfügen sowohl über verbale Kompetenzen als auch über numerische Intelligenz. Mathematikleistung ist ein zweidimensionales Konstrukt, so dass bei der Leistungsbewertung beide Dimensionen getrennt voneinander zu berücksichtigen sind."

Nachdem die Kriterien zur mathematischen Leistungsmessung identifiziert sind, folgt die Leistungsmessung im konkreten Fall durch die unterrichtenden Lehrerinnen und Lehrer. Die mathematischen Fähigkeiten der Schülerinnen und Schüler sollen durch die Lehrkräfte im Rahmen dieses Forschungsprojekts nach der geleisteten Vorarbeit auf der Grundlage der extrahierten latenten Faktoren erhoben werden.

Die Lehrerinnen und Lehrer sollen die Probanden dabei auf einer stetigen Skala von null (*sehr schlechte Fähigkeiten*) bis zehn (*herausragende Fähigkeiten*) sowohl hinsichtlich ihrer *verbalen Kompetenzen* als auch hinsichtlich ihrer *numerischen Intelligenz* bewerten. Eine solche Skala eignet sich insofern, als dass stetige Merkmale beliebige reelle Werte, meist in einem bestimmten Bereich annehmen, und haben den Vorteil, dass unter Verwendung von kombinierten Werten (z.B. Mittelwert, Median, Standardabweichung, etc.) Aussagen über bestimmte Eigenschaften der Daten getroffen werden können.

Bei der Beurteilung sollen allerdings diejenigen Aspekte vermieden bzw. zumindest abgeschwächt werden, die im Rahmen der Notengebung üblicherweise kritisiert werden. Insofern erscheint es logisch, dass die Lehrkräfte vor Abgabe ihres Urteils in einer ausführlichen Reflexionsphase hinsichtlich der Ziele des Forschungsprojekts konkret auf die Art und Hintergrundgedanken der nun vorzunehmenden Leistungsbewertung vorbereitet werden. Dabei müssen den Lehrerinnen und Lehrern sowohl die beiden zu bewertenden latenten Faktoren genaustens beschrieben als auch die Bewertung auf einer für Lehrkräfte unüblichen Bewertungsskala mit ihnen zufriedenstellend diskutiert werden. Zudem ist es dem Verfasser besonders wichtig, die Lehrkräfte darauf hinzuweisen, dass ihre Bewertungen im Unterschied zu Schulnoten keine Erziehungsziele oder andere Funktionen erfüllen müssen, so dass sie sich rein auf die Leistung der Schülerinnen und Schüler beziehen können. Darüber hinaus werden die Lehrkräfte darum gebeten, dass sie versuchen sollen, die Leistungen der Schülerinnen und Schüler möglichst nicht anhand des klassen- oder schulinternen Bezugsrahmens zu bewerten.

Um die Vergleichbarkeit der vergebenen Lehrerurteile zu steigern, werden die vorgenommenen Leistungsbewertungen am jeweiligen Klassenmittelwert zusätzlich z-standardisiert. Dies meint, dass der Klassendurchschnitt einem Wert von null entspricht. Eine positive Abweichung nach oben steht dementsprechend für eine bessere und eine negative Abweichung nach unten für eine im Vergleich zum Klassenmittelwert schlechtere Leistungen. Somit wird die Aussagekraft der Studie deutlich erhöht.

Die durchschnittliche Bewertung der *verbalen Kompetenzen* sämtlicher Probanden liegt bei **5,72** inklusive einer Standardabweichung von **2,16**. Die durchschnittliche Bewertung der *numerischen Intelligenz* der Probandengruppe beträgt **5,86** bei einer Standardabweichung von **1,97**. Dementsprechend gehen die Lehrkräfte davon aus, dass die Schülerinnen und Schüler im klassischen mathematischen Sektor etwas besser ausgebildet sind, als im sprachlichen Bereich. Beide arithmetischen Mittelwerte liegen oberhalb des Mittelpunkts der Bewertungsskala. Es kann also von einer überdurchschnittlichen Lehrerbewertung ausgegangen werden. Die Entstehung und Verteilung der Werte lässt sich allerdings durch einen Blick auf die durchschnittlichen Bewertungen der verschiedenen Schulformen erahnen.

Tabelle 7.2 stellt daher in einer 3x2-Matrix die durchschnittliche Mathematikleistung der Schülerinnen und Schüler je nach besuchter Schulform (Gymnasium, Gesamtschule, Hauptschule) und zu betrachtender mathematischer Dimension (*verbal*, *numerisch*) dar. Der Tabelle können insbesondere drei zentrale Aussagen entnommen werden.

Tabelle 7.2

Durchschnittliche Mathematikleistung der SuS

	Gymnasium	Gesamtschule	Hauptschule
verbale Kompetenzen	6,83	5,42	3,92
numerische Intelligenz	6,59	5,80	4,31

(i) Die Gymnasiasten (6,83 und 6,59) sind den Gesamt- (5,42 und 5,80) und HauptschülerInnen (3,92 und 4,31) in beiden Dimensionen deutlich überlegen. Die vermutete Leistungshierarchie zwischen den verschiedenen Schulformen findet damit Bestätigung.

(ii) Es wird deutlich, dass die Durchschnittsbewertungen der gymnasialen Probanden in beiden Dimensionen oberhalb des arithmetischen Mittelwerts sämtlicher Probanden (5,72 und 5,86) liegen. Die Leistungsfähigkeit der Gesamt- und HauptschülerInnen liegt im Mittel dagegen unterhalb des berechneten Durchschnittswerts.

(iii) Bemerkenswert ist, dass am Gymnasium die *verbalen Kompetenzen* augenscheinlich besser ausgebildet sind als die *numerische Intelligenz*. In den beiden anderen Schulformen sind die Probanden bezüglich ihrer *verbalen Kompetenzen* gleichermaßen schwächer bewertet worden (-0,38 bzw. -0,39). Diese Feststellung lässt sich *möglicherweise* durch den geringen Anteil derjenigen Gymnasiasten, die einen Migrationshintergrund aufweisen, begründen.

Zusammenfassend verdeutlicht die differenzierte Tabelle vor allem, dass die Leistungshierarchie zwischen den Schulen bei der Lehrerumfrage Bestätigung findet. Dies deutet darauf hin, dass die Lehrerinnen und Lehrer sich tatsächlich bemüht haben, die Mathematikleistung weitestgehend unabhängig von schulinternen Bezugssystemen anzugeben. Im weiteren Verlauf werden dennoch die am Klassenmittelwert z-standardisierten Bewertungen verwendet.

Dieses Vorgehen soll bei der Analyse der geschlechtsspezifischen Mathematikleistungen ein erstes Mal zum Einsatz kommen. Dazu wird jeweils das arithmetische Mittel der positiven und negativen Abweichungen der männlichen und weiblichen Probanden von ihrem jeweiligen Klassenmittelwert bestimmt.

Dem nachfolgenden Säulendiagramm ist dabei zu entnehmen, dass die männlichen Probanden in beiden mathematischen Bereichen (leicht) bessere Leistungen erbringen als die weiblichen Probanden. Bei den Beurteilungen der *verbalen Kompetenzen* machen die Lehrerinnen und Lehrer allerdings kaum einen Unterschied zwischen den Geschlechtern, so dass die Jungen im Durchschnitt verschwindend geringe 0,02 Bewertungspunkte oberhalb und die Mädchen durchschnittlich lediglich 0,02 Bewertungspunkte unterhalb des jeweiligen Klassenmittelwerts einsortiert werden. Während hinsichtlich der *verbalen Kompetenzen* die Leistungsdifferenz also kaum wahrnehmbar ist, liegen die männlichen Probanden bei der *numerischen Intelligenz* etwas deutlicher oberhalb (0,11) und die weiblichen Probanden ebenso klar unterhalb des Mittelwerts (-0,14). Damit findet, zumindest in dieser Studie, das Vorurteil Bestätigung, nachdem Jungen etwas bessere mathematische Leistungen erbringen.

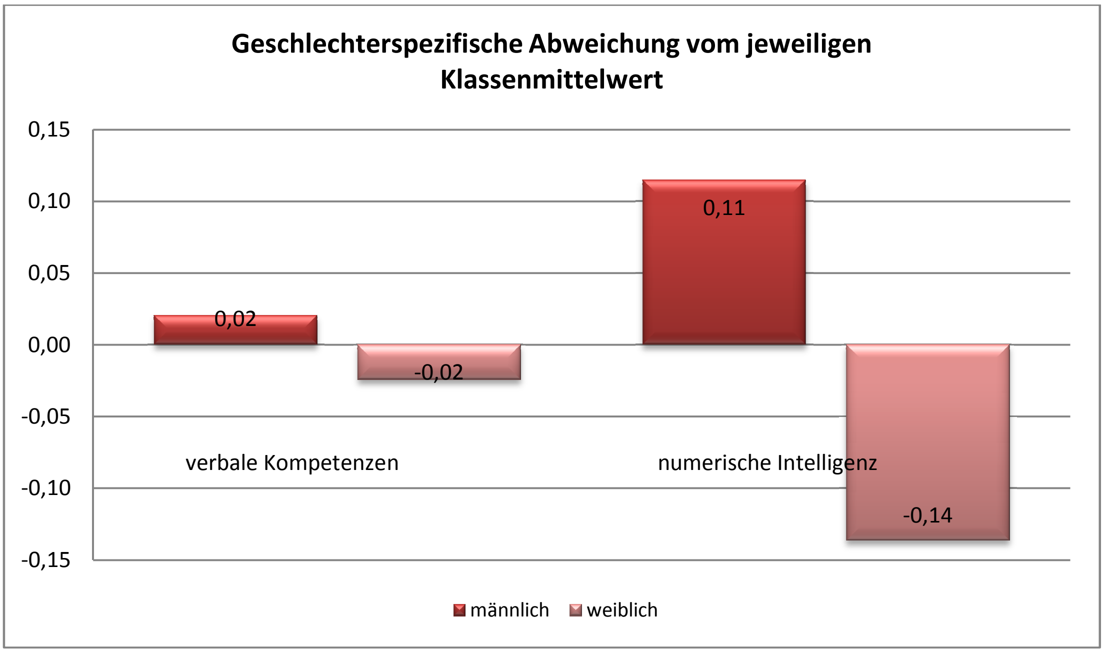


Abbildung 7.1 Geschlechtsspezifische Abweichung vom jeweiligen Klassenmittelwert

7.5 Zusammenfassung „Leistungsmessung"

Anhand verschiedener mathematischer Modelle ist die Problematik der Leistungsmessung verdeutlicht worden. Sowohl eine geeignete Auswahl von Kriterien zur Leistungsmessung als auch eine Leistungsmessung im konkreten Fall bei vorliegenden Kriterien erscheint äußerst problematisch. Dies gilt vor allem dann, wenn die Zufallsvariablen nicht sicher Intervallskalenniveau haben.

Um zu untersuchen, wie die Lehrerinnen und Lehrer im Lebensraum der Schule und im täglichen Unterricht Mathematikleistung definieren und messen, werden daher zwei Fragebogenaktionen durchgeführt und mit Hilfe der modifizierten Faktorenanalyse ausgewertet. Dabei kann der Verfasser feststellen, dass die Mathematikleistung für die teilnehmenden Lehrerkräfte aus zwei Faktoren besteht: der *verbalen Kompetenz* und der *numerischen Intelligenz*. Sämtliche Schülerinnen und Schüler werden daraufhin auf einer stetigen Bewertungsskala zwischen null (*sehr schlecht*) und zehn (*sehr gut*), möglichst schulformübergreifend bewertet. Um die Vergleichbarkeit der Lehrerurteile zu erhöhen, werden die Bewertungen anhand des jeweiligen Klassenmittelwerts z-standardisiert. Dabei zeigt sich, dass die Gymnasiasten vor den Gesamt- und vor den HauptschülerInnen am Besten bewertet werden. Zudem lässt sich feststellen, dass die Schülerinnen und Schüler an Gesamt- und Hauptschulen über bessere *numerische* als *verbale Kompetenzen* verfügen, während es am Gymnasium genau umgekehrt ist. Zudem kann gezeigt werden, dass Jungen in beiden Dimensionen die etwas besseren Leistungen erbringen als Mädchen.

8. Interdependenzen zwischen Fernsehverhalten und Mathematikleistung: die isolierte Auswertung

8.1 Die Methode der isolierten Auswertung

In den bisherigen methodischen Kapiteln, die sich auf das Forschungsprojekt dieser Arbeit bezogen haben, wurden die tägliche Fernsehdauer, die latenten Faktoren des Fernsehkonsums, das Freizeitverhalten sowie die Mathematikleistungen von 232 Probanden aus drei verschiedenen Schulformen erhoben. Damit ist die Vorarbeit für die finale Auswertung abgeschlossen. Das Fernsehverhalten und die mathematischen Kompetenzen können in Bezug gesetzt werden. Dazu wird wie folgt vorgegangen: Zunächst werden die drei Variablen getrennt voneinander mit der Mathematikleistung der Probanden verglichen. Es wird dann möglich sein für jeden Risikofaktor explizit anzugeben, ob ein Zusammenhang zwischen dem Lehrerurteil der mathematischen Fähigkeiten und dem jeweiligen Fernseh- oder Freizeitverhalten besteht. Nach diesem ersten Schritt werden die drei erhobenen Faktoren auf alle möglichen Arten miteinander kombiniert, so dass letztlich konkret angegeben wird, wie gut oder schlecht die Mathematikleistungen von Jugendlichen in Abhängigkeit von der Fernsehdauer, der Fernsehart und dem sonstigen Freizeitverhalten, jeweils gemessen am Klassenmittelwert, sind.

8.2 Fernsehdauer und Mathematikleistung

Zunächst untersuche man den möglichen Zusammenhang zwischen der täglichen Fernsehdauer und den entsprechenden Mathematikleistungen der Probanden. Zu diesem Zweck sind die Probanden im Verlauf des Abschnitts „Dauer und Häufigkeit des Fernsehkonsums” mit Hilfe des Ward-Algorithmus entsprechend ihrer durchschnittlichen täglichen Fernsehdauer in die vier Gruppen *Wenigseher*, *Mittelseher*, *Vielseher* und *Exzessivseher* eingeteilt worden. Nun kann die Fernsehdauer der jeweils zu einer Gruppe zusammengefassten Jugendlichen mit den gemessenen Mathematikleistungen in Bezug gesetzt werden. Das Vorgehen zur Messung der

Mathematikleistungen wurde bereits im siebten Kapitel ausführlich beschrieben. Zusammenfassend bleibt lediglich zu wiederholen, dass die Lehrerinnen und Lehrer für jeden Probanden auf einer stetigen Skala Bewertungspunkte zwischen 0 (*sehr schlecht*) und 10 (*sehr gut*) für die zwei identifizierten Dimensionen *numerische Intelligenz* und *verbale Kompetenzen* vergeben haben. Um die Vergleichbarkeit der verschiedenen Lehrerurteile zu steigern, sind die Bewertungen am Klassenmittelwert z-standardisiert worden.

Die Abbildung 8.1 präsentiert den fundamentalen Zusammenhang zwischen den beiden Variablen:

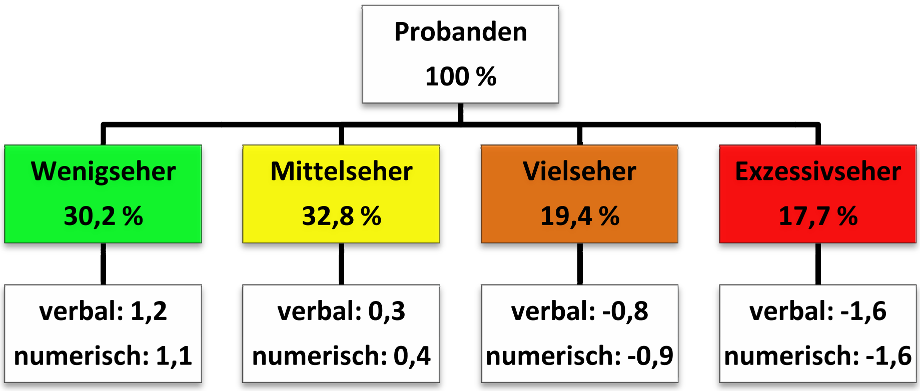


Abbildung 8.1 Fernsehdauer und Mathematikleistungen

Das Dendrogramm zeigt zunächst erneut, wie sich sämtliche Probanden auf die vier verschiedenen Gruppen ähnlicher Fernsehdauer verteilen. Die beiden größten Cluster bilden die *Wenig-* (30,2 %) und die *Mittelseher* (32,8 %). 19,4 % der jugendlichen Probanden gehören zu den *Vielsehern*. Der risikobehafteten Gruppe der *Exzessivseher* gehören 17,7 % der UntersuchungsteilnehmerInnen an.

Folgt man dem Pfad weiter, lassen sich die Lehrerurteile als positive oder negative durchschnittliche Abweichung vom jeweiligen Klassenmittelwert für die *verbalen Kompetenzen* („verbal") und die *numerische Intelligenz* („*numerisch*") ablesen.

Dem Dendrogramm darf allerdings keine Einflussrichtung entnommen werden. Ob die Fernsehdauer zu den entsprechenden Lehrerurteilen führt, oder ob es umgekehrt ist, darf nicht auf Grundlage der Leserichtung in das Dendrogramm hineininterpretiert werden. Es wird lediglich ein Zusammenhang präsentiert!

Zunächst betrachte man die *Wenigseher*, die in beiden gemessenen Dimensionen mehr als deutlich oberhalb des jeweiligen Klassenmittelwerts liegen. Im Durchschnitt sind diese Probanden bis zu **1,2** Bewertungspunkte besser als der Vergleichswert. Dieses Ergebnis legt die Interpretation nahe, dass eine stark begrenzte tägliche Fernsehdauer in einem engen positiven Zusammenhang zur Entwicklung der Mathematikleistungen steht. Bereits die mathematischen Leistungen der *Mittelseher* werden deutlich schlechter eingestuft. Dennoch sortieren sich auch diese Jugendlichen noch knapp, maximal **0,4** Bewertungspunkte, oberhalb des Klassendurchschnitts ein, büßen aber damit im Vergleich zu den *Wenigsehern* einen Großteil des positiven Zusammenhangs ein. Nichtsdestotrotz steht der Fernsehkonsum von höchstens 2 Stunden 15 Minuten in einem überdurchschnittlichen Zusammenhang zu den Bewertungen. Diese Beobachtung gilt allerdings bei der nächsten Gruppe nicht mehr. Die Mathematikleistungen der *Vielseher* werden durchschnittlich mindestens **0,8** Bewertungspunkte schlechter als der Klassenmittelwert beurteilt. Ein zeitlich erhöhter Fernsehkonsum und die erhobenen Rechenleistungen befinden sich demnach in einem negativen Zusammenhang. Im Mittel schneiden die *Vielseher* bei der Lehrerumfrage **2,0** Bewertungspunkte schlechter ab als die *Wenigseher*. Getoppt wird dieses Ergebnis lediglich von den *Exzessivsehern*, die über 3 Stunden 24 Minuten fernsehen und eine um **1,6** Punkte unterhalb des Vergleichswerts liegende Durchschnittsbewertung erhalten.

Das Säulendiagramm in Abbildung 8.2 verdeutlicht dieses Ergebnis grafisch. Die x-Achse ist jeweils für die Gruppen der Fernsehdauer vorgesehen, während die y-Achse die durchschnittliche Abweichung wiedergibt. Die blauen Säulen präsentieren die Ergebnisse für die *verbalen Kompetenzen*, die roten Säulen die Ergebnisse der *numerischen Intelligenz*.

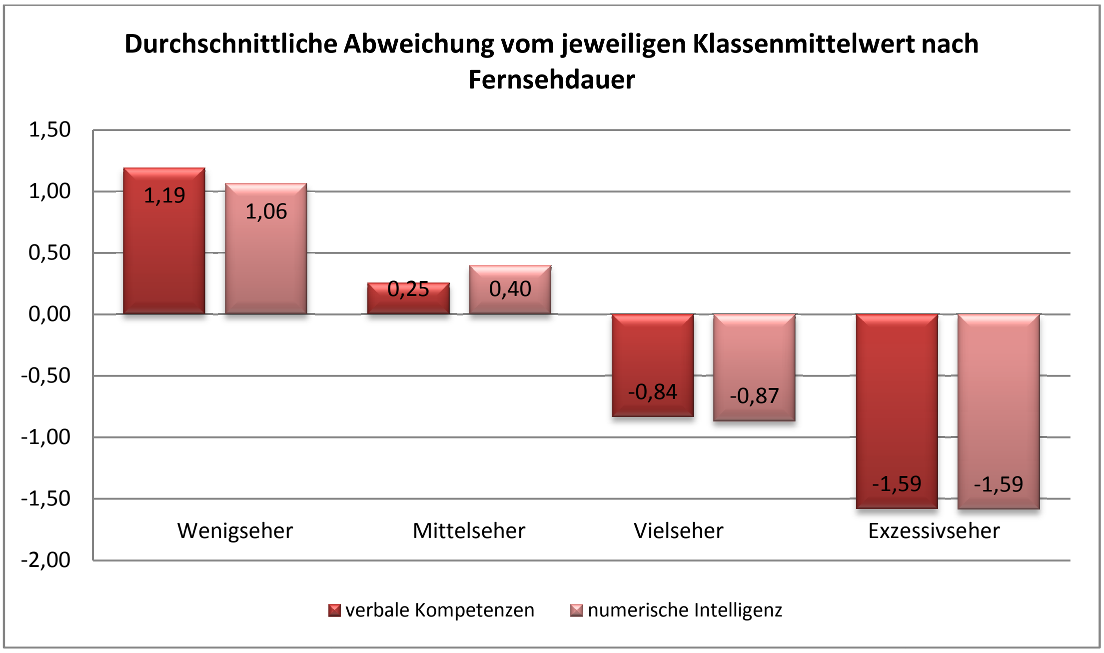


Abbildung 8.2 Durchschnittliche Abweichung vom jeweiligen Klassenmittelwert nach Fernsehdauer

Da der den möglichen Zusammenhang beschreibende arithmetische Mittelwert nicht zu den robusten Lagemaßen zu zählen ist, besteht das Risiko, dass etwaige Ausreißer das Ergebnis verzerren. Um die Gefahr einer solchen Fehlinterpretation hier zu minimieren, wird zusätzlich der prozentuale Anteil derjenigen Probanden angegeben, die in Abhängigkeit von ihrer Fernsehdauer eine überdurchschnittliche Lehrerbeurteilung erhalten haben.

Das Ergebnis kann an Abbildung 8.3 abgelesen werden und bestätigt die bisherigen Beobachtungen in vollem Umfang:

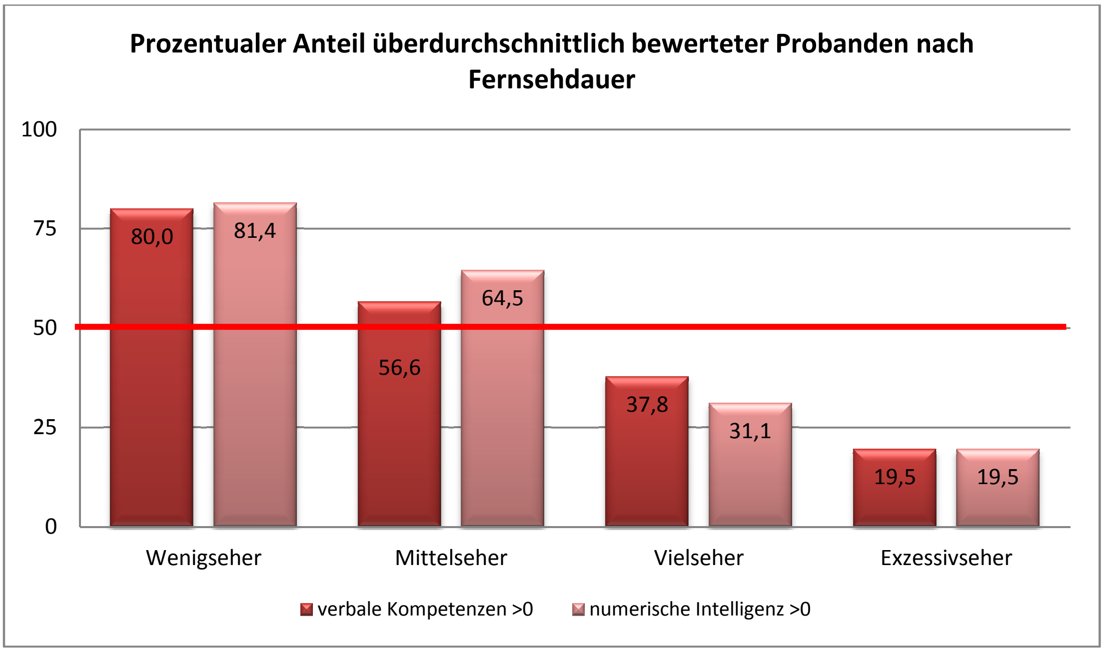


Abbildung 8.3 Prozentualer Anteil überdurchschnittlich bewerteter Probanden nach Fernsehdauer

Etwa 80 % der *Wenigseher* erhalten in den beiden mathematischen Dimensionen eine überdurchschnittliche Bewertung. Dieser Wert bekräftigt die Vermutung, dass eine geringe tägliche Fernsehdauer mit positiven mathematischen Leistungen von Jugendlichen zusammenhängt. Bei den *Mittelsehern* befinden sich lediglich 56,6 % (*verbal*) und 64,5 % (*numerisch*) im überdurchschnittlichen Bereich. Der beobachtete Leistungsabfall zu den *Wenigsehern* wird in dieser Statistik ebenfalls sichtbar. Nichtsdestotrotz liegen beide *Mittelseher*-Prozentangagen noch oberhalb der im Diagramm rot markierten magischen 50 Prozentgrenze. Dies meint, dass mehr als die Hälfte aller *Mittelseher* eine im Vergleich positive Bewertung erhalten haben. Von solchen Werten sind die *Viel*- und vor allem die *Exzessivseher* weit entfernt, liegt der prozentuale Anteil der überdurchschnittlich gut bewerteten Jugendlichen bereits bei den *Vielsehern* nur noch bei 37,8 % (*verbal*) bzw. 31,1 % (*numerisch*), wobei ein offensichtlicher Vorteil der *verbalen Kompetenzen* gegenüber der *numerischen Intelligenz* identifiziert werden kann. Da das Fernsehen auch ein akustisches Medium ist, überrascht dieser Unterschied nicht. Bei den *Exzessivsehern* fällt dieser Wert noch weiter ab. Nur knapp ein Fünftel dieser Probanden erhält eine positive Beurteilung seiner mathematischen Leistungen; oder anders formuliert: Die Lehrerinnen und Lehrer assoziieren 80 % (!) der *Exzessivseher* mit einer unterdurchschnittlichen Mathematikleistung.

Als erstes Zwischenergebnis der Auswertung lässt sich ein offensichtlich negativer Zusammenhang zwischen Fernsehdauer und Mathematikleistung feststellen. Die *Wenigseher* erhalten vor den *Mittel-* und den *Vielsehern* die mit Abstand besten Lehrerbewertungen. Die *Exzessivseher* landen abgeschlagen auf dem letzten Rang.

Um das Zustandekommen dieses Ergebnisses detaillierter zu untersuchen, sollen nach diesem ersten groben Überblick im Folgenden zusätzlich das Geschlecht und die Schulform der Probanden mit einbezogen werden.

Tabelle 8.1

Abweichung vom (geschlechtsspezifischen) Klassenmittelwert nach Fernsehdauer und Geschlecht

	verbale Kompetenz		numerische Kompetenz	
	weiblich	männlich	weiblich	männlich
Wenigseher	1,10	**1,24**	1,06	1,01
Mittelseher	0,24	0,20	0,33	0,32
Vielseher	-0,77	-0,79	-0,80	-0,80
Exzessivseher	-1,38	**-1,69**	-1,43	**-1,55**

Tabelle 8.1 zeigt die Abweichungen vom (geschlechtsspezifischen) Klassendurchschnitt bezüglich der *verbalen Kompetenz* und der *numerischen Intelligenz* nach Fernsehdauer und Geschlecht. Zwar bestätigen die Daten für beide Geschlechter die Tendenz, dass die *Wenigseher* in beiden mathematischen Dimensionen am Besten abschneiden und die Leistungsfähigkeit der weiteren Gruppen mit steigender Fernsehdauer stetig abnimmt, doch offenbaren sich dennoch zwei offensichtliche Unterschiede zwischen Mädchen und Jungen. Zum einen werden vor allem die *verbalen Kompetenzen* der männlichen Probanden mit einer geringen Fernsehdauer (**1,24**) sehr gut bewertet. Zum anderen weichen die Leistungen der *exzessiv* fernsehenden Jungen in beiden mathematischen Bereichen deutlich stärker nach unten ab, als dies bei den Mädchen der Fall ist. Damit werden die geschlechtsspezifischen Forschungsergebnisse des Kriminologischen Forschungsinstituts Niedersachsens bestätigt: Vor allem die Fähigkeiten derjenigen Jungen, die ein fragwürdiges Fernsehverhalten zeigen, sind besonders schwach entwickelt, während die männlichen Jugendlichen mit einer geringen Fernsehdauer zumindest im sprachlichen Sektor besser abschneiden als der Rest der Stichprobe!

Zur Absicherung dieses Ergebnisses folgt mit den Abbildungen 8.4 und 8.5 erneut die prozentuale Verteilung der überdurchschnittlich und unterdurchschnittlich bewerteten Jungen und Mädchen in Abhängigkeit von der individuellen Fernsehdauer. Dabei soll konkret das Zustandekommen der beiden identifizierten geschlechtsspezifischen Unterschiede erörtert werden:

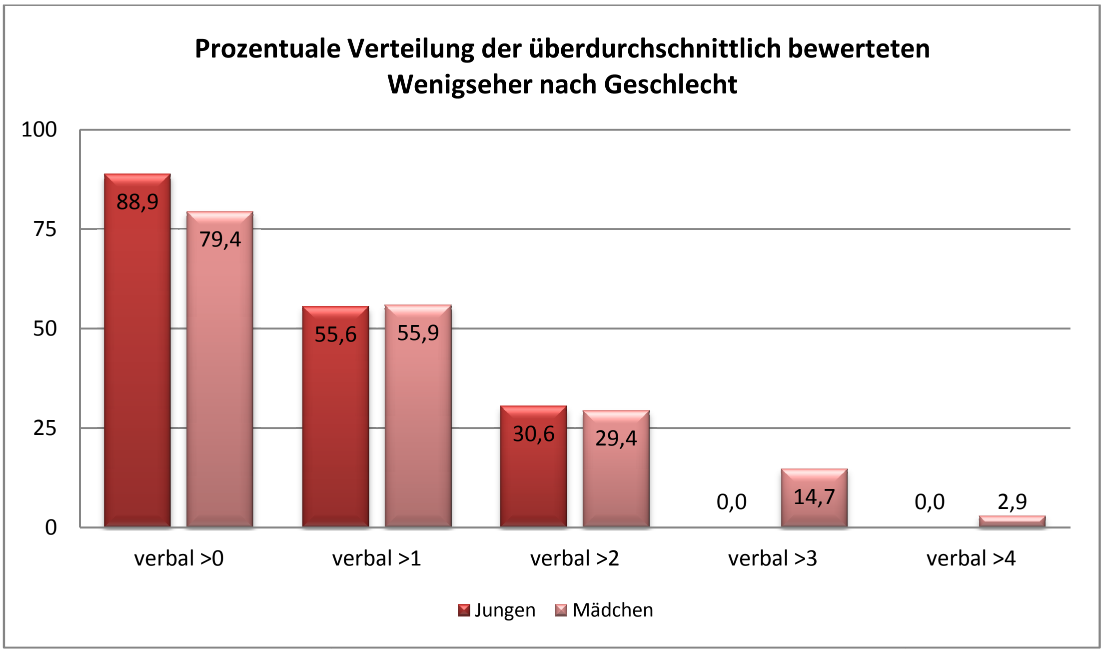


Abbildung 8.4 prozentuale Verteilung der überdurchschnittlich bewerteten Wenigseher nach Geschlecht

Die Abbildung 8.4 setzt sich zunächst mit der prozentualen Verteilung der überdurchschnittlich bewerteten *Wenigseher* nach Geschlecht auseinander. Dies geschieht aufgrund des identifizierten Unterschieds zwischen Jungen und Mädchen anhand der *verbalen Kompetenzen*. Die Visualisierung der Daten bestätigt die Vormachtstellung der männlichen Probanden mit einer zeitlich stark limitierten Fernsehdauer im *verbalen* Sektor. Fast 90 % der Jungen erhalten eine überdurchschnittliche Bewertung. Bei den Mädchen sind es nur knapp 80 %. Nahezu identisch ist die Verteilung bei denjenigen Probanden, die mindestens einen (56 %) bzw. zwei (30 %) Bewertungspunkte über dem Durchschnitt liegen. Lediglich bei den Mädchen kommt es in der Folge zu Ausreißern nach oben. 15 % werden über 3 Bewertungspunkte besser bewertet, 3 % sogar um über 4 Punkte.

Die zweite geschlechtsspezifische Unterscheidung, nämlich dass die mathematischen Kompetenzen von Jungen bei einem *exzessiven* Fernsehkonsum schlechter als die Fähigkeiten der Mädchen beurteilt werden, wird mit Hilfe von dem Diagramm in Abbildung 8.5 näher untersucht. Dies geschieht erneut anhand der *verbalen Kompetenzen*, da die identifizierte Leistungsdifferenz hier besonders deutlich wird. Für die *numerische Intelligenz* gelten die Ausführungen analog.

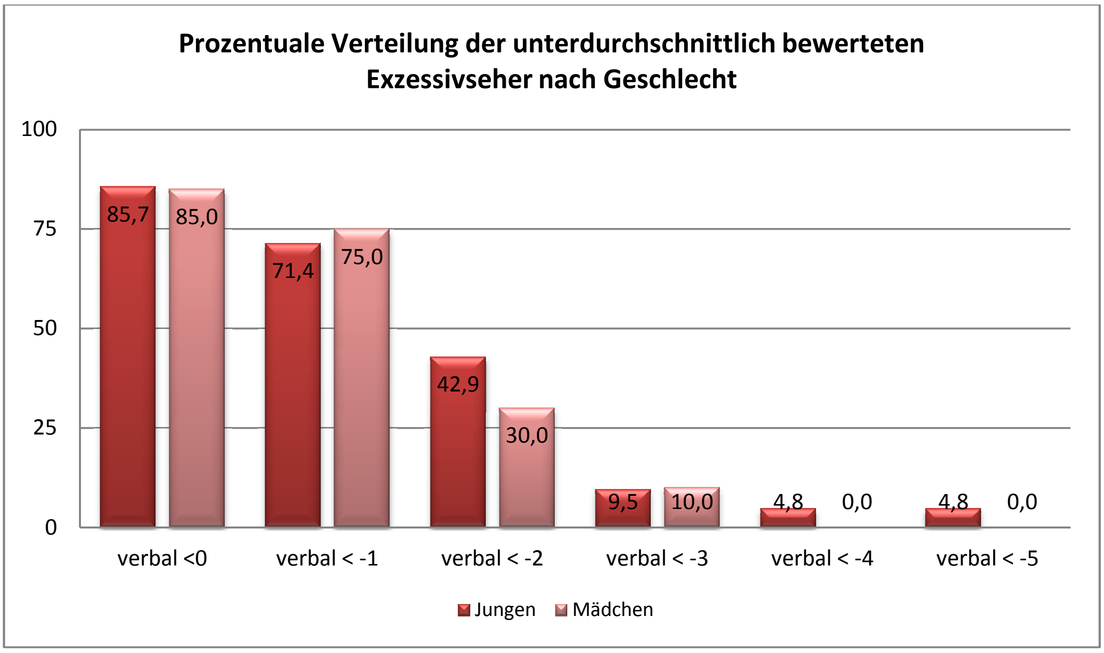


Abbildung 8.5 Prozentuale Verteilung der unterdurchschnittlich bewerteten Exzessivseher nach Geschlecht

Auf den ersten Blick ist nicht zu erkennen, warum die männlichen Probanden stärker nach unten abweichen als die weiblichen. Sowohl 85,7 % der Jungen als auch 85 % der Mädchen erhalten eine unterdurchschnittliche Bewertung. Hier liegen keine Unterschiede vor, welche die Leistungsdifferenz rechtfertigen. Der stark negative Zusammenhang zwischen mathematischen Fähigkeiten einerseits und *exzessivem* Fernsehkonsum andererseits bei den Jungen ist dementsprechend weniger einer im Vergleich zu den Mädchen besonders hohen Anzahl unterdurchschnittlicher Bewertungen geschuldet. Vielmehr ist er Folge zahlreicher extrem schwacher Lehrerurteile, wie die weitere Betrachtung des Diagramms verrät. Es fällt auf, dass wesentlich mehr männliche Jugendliche eine um über 2 Bewertungspunkte nach unten abweichende Lehrerbeurteilung erhalten haben. Dies gilt für 43 % der *exzessiv* fernsehenden Jungen, betrifft aber lediglich 30 % der Mädchen. Einige männliche Untersuchungsteilnehmer mit *exzessivem* Fernsehkonsum erhalten sogar Bewertungen, die über 5 Punkte vom geschlechtsspezifischen Klassenmittelwert abweichen (5 %). Diese Bündelung an besonders negativen Beurteilungen begründet letztlich, warum die *exzessiv* sehenden Jungen noch schlechter abschneiden als die im Fernsehverhalten vergleichbaren Mädchen!

Nach der Analyse möglicher geschlechtsspezifischer Unterschiede bezüglich des Einflusses der Fernsehdauer auf die Mathematikleistung von Jugendlichen steht nun eine schulformabhängige Diskussion im Mittelpunkt:

Tabelle 8.2 gibt dazu nachfolgend die Abweichungen vom Klassenmittelwert in Abhängigkeit von der täglichen Fernsehdauer und der besuchten Schulform, wiederum auf beide mathematischen Dimensionen bezogen, an.

Tabelle 8.2

Abweichung vom Klassenmittelwert nach Fernsehdauer und Schulform

	verbale Kompetenz			numerische Intelligenz		
	Gy	Ge	Hs	Gy	Ge	Hs
Wenigseher	0,84	1,64	**2,73**	0,65	1,55	**2,95**
Mittelseher	**-0,26**	0,34	**1,87**	**-0,23**	0,62	**1,73**
Vielseher	**-1,96**	-0,48	-0,72	**-1,70**	-0,68	-0,64
Exzessivseher	**-2,24**	-1,59	-1,30	**-1,34**	-1,81	-1,37

An allen drei Schulformen wird die grundsätzliche Tendenz (bis auf eine Ausnahme) bestätigt. Dennoch fallen zwei Aspekte ins Auge. Einerseits weichen die mathematischen Kompetenzen der *Wenigseher* der Hauptschule (Hs) mit Abstand am stärksten nach oben ab. Die Leistungen dieser Jugendlichen liegen **2,73** (*verbal*) bzw. **2,95** (*numerisch*) Punkte über der hauptschulinternen Durchschnittsbewertung. Selbst die *Mittelseher* erreichen noch beeindruckende Zahlen, schließlich ist ihre Abweichung nach oben höher als bei den *Wenigsehern* der anderen beiden Schulformen. Scheinbar hängen bei Hauptschülerinnen und -schülern gute Mathematikleistungen besonders stark mit einem zeitlich geregeltem Fernsehkonsum zusammen. Leider besteht die Hauptschule in dieser Altersklasse zu zwei Dritteln aus *viel-* oder *exzessiv* fernsehenden Jugendlichen, so dass die positiven Zusammenhänge leider bei nur wenigen Probanden zu beobachten sind. Andererseits fällt auf, dass die *Viel-* und *Exzessivseher* des Gymnasiums im Vergleich zu ihren weniger fernsehenden Klassenkameraden besonders schlecht abschneiden. Schon die *Mittelseher* liegen unterhalb des Vergleichswerts. Lediglich die Fernsehdauer der *Wenigseher* steht in einem positiven Zusammenhang zu den mathematischen Leistungen, der im schulformübergreifenden Vergleich dennoch als eher gering einzustufen ist. Die schulischen Anforderungen des Gymnasiums sind *möglicherweise* bei erhöhtem Fernsehkonsum nicht zur Zufriedenheit zu bewältigen. Im Unterschied zur Hauptschule lässt sich glücklicherweise festhalten, dass das Gymnasium zu über 50 % aus *Wenigsehern* besteht. Lediglich 16 % der Gymnasiasten schauen *viel* oder *exzessiv* fern. Die Ergebnisse der Gesamtschule sortieren sich insgesamt zwischen den vorgestellten Schulformen ein.

Die Abbildung 8.6 reflektiert das Zustandekommen der obigen Ergebnisse und bekräftigt diese, indem für sämtliche Schulformen und beide mathematischen Faktoren der Anteil überdurchschnittlicher Lehrerurteile in Abhängigkeit von der Fernsehdauer illustriert wird:

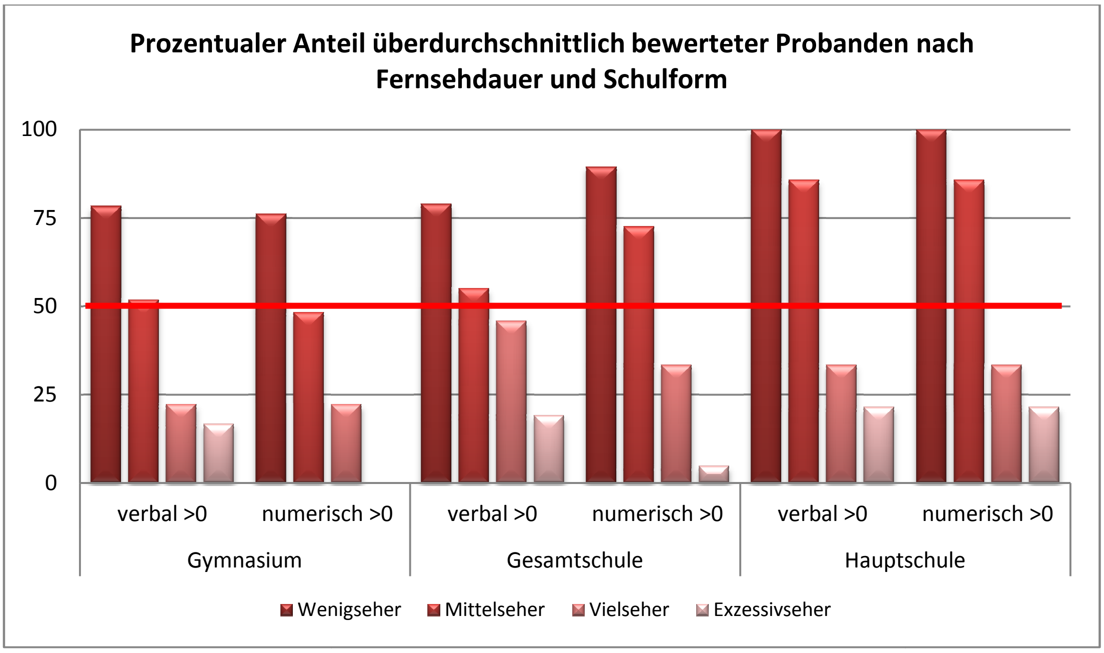


Abbildung 8.6 Prozentualer Anteil überdurchschnittlich bewerteter Probanden nach Fernsehdauer und Schulform

Nur ein vergleichsweise geringer Anteil (77 %) der gymnasialen *Wenigseher* wird in den beiden mathematischen Bereichen positiv bewertet. Sowohl an der Gesamtschule als auch an der Hauptschule gilt dies für wesentlich mehr Jugendliche. Die herausragende Abweichung der *wenig* und *mittellang* fernsehenden Hauptschülerinnen und -schüler wird ebenfalls durch das Diagramm erklärt. Alle *Wenigseher* erhalten eine überdurchschnittliche Bewertung. Dies gilt ebenfalls für über 85 % der *Mittelseher*, aber gerade mal für ein Drittel der *Vielseher* und 21 % der *Exzessivseher*. Das Leistungsgefälle zwischen dem gemäßigten und dem überhöhten Fernsehkonsum wird hier besonders deutlich. Gemessen an der durchschnittlichen Abweichung vom Klassenmittelwert gilt die *numerische Intelligenz* der *Exzessivseher* als besser ausgebildet als bei den *Vielsehern*. Der fehlende Balken in diesem Diagramm relativiert aber selbst diese Ausnahme. Kein *exzessiv* fernsehender Gymnasiast erhält eine überdurchschnittliche Bewertung.

Zusammengefasst lässt sich eine eindeutig negative Beziehung zwischen Fernsehdauer und Mathematikleistung festhalten. Eine zunehmende Fernsehdauer ist mit nachlassenden *verbalen* und *numerischen* Fähigkeiten verbunden. Insgesamt bestätigt sich die Vermutung über alle Schulformen und beide Geschlechter hinweg.

8.3 Latente Faktoren des Fernsehkonsums und Mathematikleistung

Nach dem identifizierten Zusammenhang zwischen der Fernsehdauer einerseits und der Mathematikleistung der Probanden andererseits wird nun diskutiert, inwieweit die individuelle Gestaltung des Fernsehprogramms in ähnlicher Weise mit den Lehrerurteilen verbunden ist. Aus diesem Grund sind die Probanden im Verlauf des Abschnitts „Die latenten Faktoren des

Fernsehkonsums" mit Hilfe der Faktoren- und Clusteranalyse gemäß ihres Antwortverhaltens in fünf, ihrem allgemeinen Fernsehverhalten am ehesten entsprechende Gruppen eingeteilt worden. Dabei handelt es sich um die Gruppen *gemischte Unterhaltung*, *Wettkampf und Bildung*, *niveauarme Unterhaltung*, *humorvolle Unterhaltung* oder *musikalische Unterhaltung*. Nun können die latenten Faktoren des Fernsehkonsums in analoger Weise mit den gemessenen Mathematikleistungen in Bezug gesetzt werden.

Das Dendrogramm in Abbildung 8.7 stellt das Ergebnis grafisch dar, wobei wiederum lediglich Zusammenhänge zu beobachten sind:

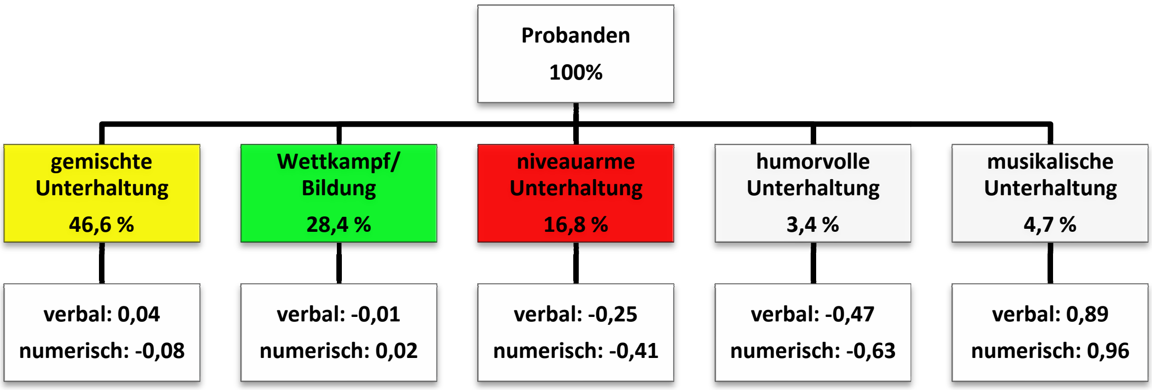


Abbildung 8.7 Latente Faktoren des Fernsehkonsums und Mathematikleistungen

Ähnlich wie das erste, die Fernsehdauer betreffende Dendrogramm erinnert auch Version 8.7 zunächst noch einmal an die Verteilung der Probanden auf die fünf latenten Faktoren des Fernsehkonsums. Ein Großteil (46,6 %) hat kein durch spezielle Vorlieben geprägtes Fernsehprogramm und bevorzugt *gemischte Unterhaltung*. Knapp 30 % verfolgen am Liebsten verschiedene *Wettkämpfe und Bildungssendungen*, während bei etwa 16 % der Probanden *niveauarme* Doku-Soaps und Castingshows auf der Tagesordnung stehen. *Humorvolle* oder *musikalische* Sendungen sind zwar beliebt, werden aber von kaum einem Probanden überwiegend geschaut. Die Anzahlen sind so gering, dass sich keinerlei statistisch abgesicherten Aussagen über einen möglichen Zusammenhang zwischen diesen beiden Genres und den Mathematikleistungen treffen lassen, so dass sie im weiteren Verlauf kaum Berücksichtigung finden werden.

Analog zur ersten Auswertungsrunde können nun die Lehrerurteile jeweils als positive oder negative Abweichung vom jeweiligen Klassenmittelwert für die *verbalen Kompetenzen* und die *numerische Intelligenz* abgelesen werden, wobei das Hauptaugenmerk auf die ersten drei, statistisch relevanten, Faktoren gelegt werden sollte.

Bereits auf den ersten Blick fallen die im Vergleich zur Fernsehdauer wesentlich geringeren mathematischen Leistungsabweichungen auf. Vor allem die *gemischte Unterhaltung* und der Faktor *Wettkampf und Bildung* scheinen so gut wie keinen messbaren Zusammenhang mit den Mathematikleistungen zu besitzen. Die Abweichungen befinden sich lediglich im kaum wahrnehmbaren Hundertstelbereich. Während dies bei der *gemischten Unterhaltung*, die sowohl aus anspruchsvollen als auch aus anspruchsarmen Fernsehsendungen besteht, durchaus eine erwartbare Erkenntnis ist, wird das Ergebnis Befürworter von *Wettkampfs*- und *Bildungssendungen* sicherlich enttäuschen. Selbst ein niveauvoller Fernsehkonsum ist, zumindest bezogen auf die Mathematikleistungen von Jugendlichen, nicht mit nennenswerten Verbesserungen im Vergleich zu Schülerinnen und Schülern mit einem anderen bevorzugten Fernsehprogramm verbunden. Lediglich diejenigen Probanden, die am liebsten *niveauarmes* Fernsehen präferieren, weichen auffälliger und negativ vom jeweiligen Klassenmittelwert ab. Die *verbalen Kompetenzen* liegen **0,25** und die *numerische Intelligenz* sogar **0,41** Bewertungspunkte unterhalb des Durchschnitts, wobei auch dieser Zusammenhang geringer ausfällt als befürchtet. Es lässt sich schlussfolgern, dass die mathematischen Leistungen von Jugendlichen mit den latenten Faktoren des Fernsehverhaltens kaum im Zusammenhang stehen, es sei denn, es wird *niveauarmes* Fernsehen geschaut. Die deutlichen Abweichungen nach unten bei der *humorvollen Unterhaltung* und nach oben bei der *musikalischen Unterhaltung* sind durchaus überraschend, aber nicht repräsentativ.

Die Abbildung 8.8 verdeutlicht das Ergebnis für die ersten drei Faktoren nach dem bekannten Strickmuster grafisch:

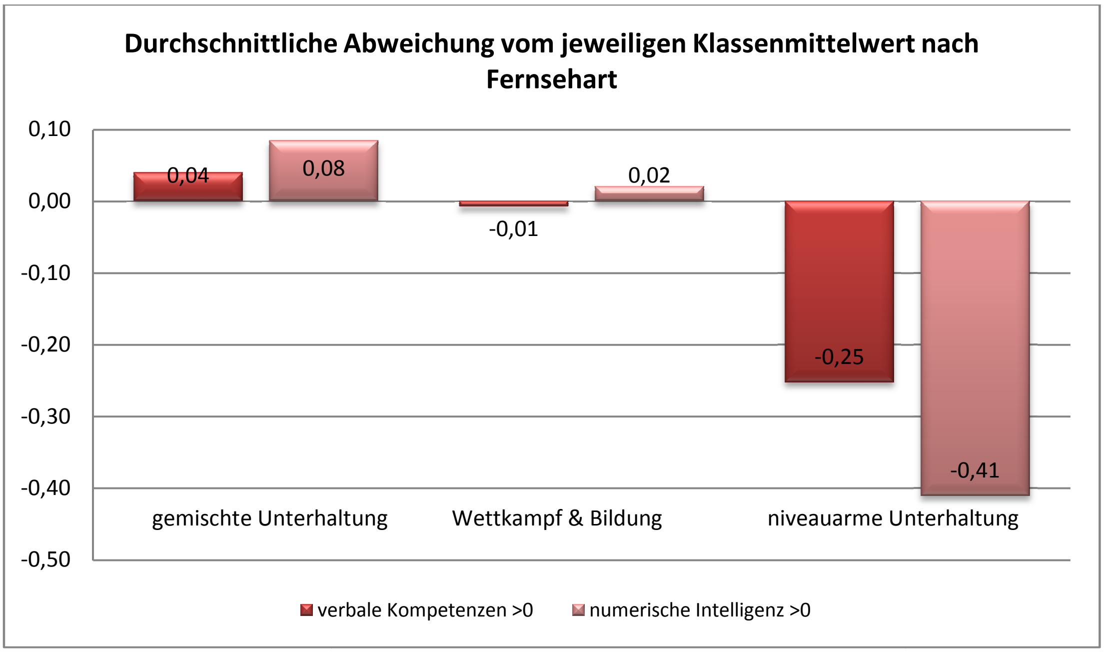


Abbildung 8.8 Durchschnittliche Abweichung vom jeweiligen Klassenmittelwert nach Fernsehart

In Anlehnung an das Vorgehen bei der Analyse der Fernsehdauer soll im Folgenden zusätzlich der prozentuale Anteil überdurchschnittlich bewerteter Probanden für die drei bedeutenden latenten Faktoren angegeben werden, um einerseits das Risiko einer Fehlinterpretation zu senken und andererseits detailliertere Einblicke zu erhalten. Dies geschieht in Abbildung 8.9:

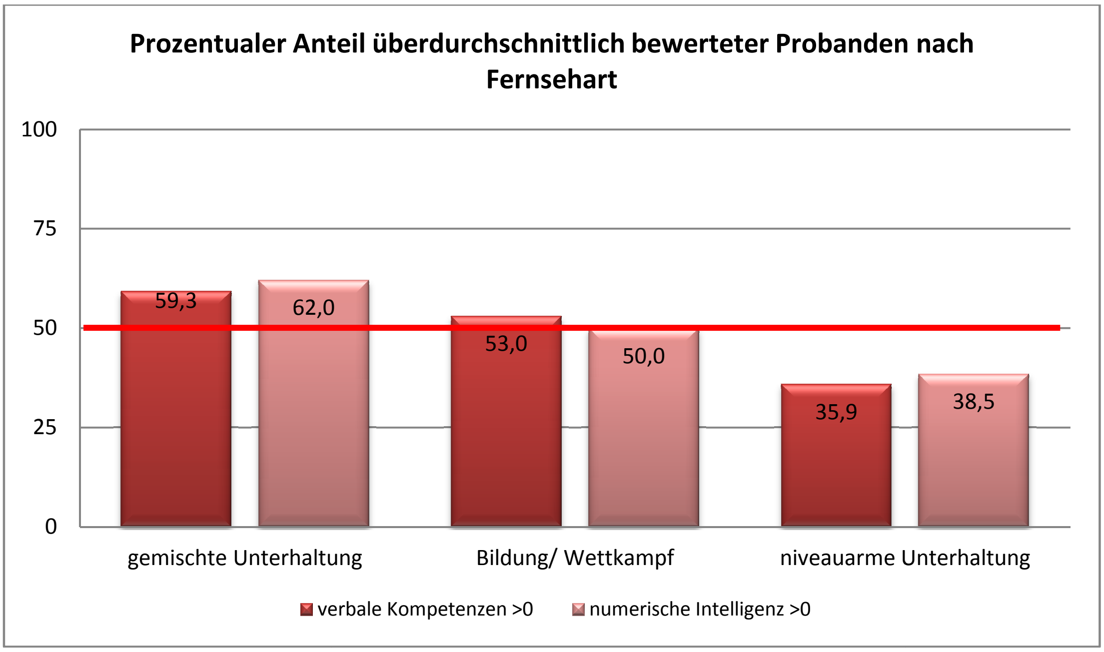


Abbildung 8.9 Prozentualer Anteil überdurchschnittlich bewerteter Probanden nach Fernsehart

Insgesamt bestätigen die in dem Diagramm abgebildeten Daten das bisherige Ergebnis vor allem für den zweiten und den dritten latenten Faktor. Jugendliche, die vorwiegend *Bildungs- und Wettkampfsendungen* schauen, erhalten nicht nur im Mittel durchschnittliche Lehrerbeurteilungen, sie werden auch exakt zur Hälfte besser bzw. schlechter als der jeweilige Klassenmittelwert bewertet. Dieser Faktor scheint tatsächlich weder in einem positiven noch in einem negativen Zusammenhang zu den Mathematikleistungen zu stehen. Probanden, die *niveauarme* Fernsehsendungen bevorzugen, erhalten nur selten eine überdurchschnittliche Lehrerbeurteilung. Die Tendenz eines negativen Zusammenhangs zwischen dieser Fernsehart und den Mathematikleistungen verhärtet sich dementsprechend. Lediglich bei der *gemischten Unterhaltung* zeigt sich ein etwas anderes Bild, als die durchschnittlichen Abweichungen, in Abbildung 8.7 dargestellt, vermuten lassen. Insgesamt erhalten nämlich etwa 60 % dieser Probanden eine nach oben abweichende Beurteilung, wobei ein Großteil der negativen Bewertungen deutlich (mindestens einen Bewertungspunkt) unterhalb des jeweiligen Klassenmittelwerts liegt. Die positiven Bewertungen weichen im Gegensatz dazu in vielen Fällen nur leicht nach oben ab.

Als zweites Zwischenergebnis der Auswertung lässt sich folgerichtig festhalten, dass zwar ein negativer Zusammenhang zwischen *niveauarmer Fernsehkost* und Mathematikleistungen besteht, dieser aber relativ gering ausfällt. Bildungssendungen lassen dagegen weder einen positiven noch einen negativen Zusammenhang zu den mathematischen Kompetenzen der jugendlichen Probanden vermuten.

Im Sinne einer detaillierten Auswertung der erhobenen Daten wird dieser Gesamteindruck um Einblicke in die geschlechts- und schulformspezifischen Untersuchungsergebnisse ergänzt. Den Anfang macht dabei die Unterscheidung nach Geschlecht, abgebildet in Tabelle 8.3:

Tabelle 8.3

Abweichung vom (geschlechtsspezifischen) Klassenmittelwert nach Fernsehart und Geschlecht

	verbale Kompetenz		numerische Kompetenz	
	weiblich	männlich	weiblich	männlich
gemischte Unterhaltung	**0,18**	-0,19	-0,04	-0,10
Wettkampf und Bildung	-0,09	**0,40**	0,07	**0,28**
niveauarme Unterhaltung	-0,09	-0,02	**-0,11**	**-0,24**

Wenngleich die Abweichungen vom jeweiligen (geschlechtsspezifischen) Klassenmittelwert auch bei der Differenzierung in Jungen und Mädchen grundsätzlich für beide Geschlechter gering bleiben, lassen sich dennoch auffällige Unterschiede erkennen. Demnach steht die Art des bevorzugten Fernsehprogramms bei den männlichen Untersuchungsteilnehmern in einem wesentlich größeren Zusammenhang zu den Leistungen in Mathematik als bei den weiblichen Probanden. Die Mädchen weichen sendungsunabhängig nur schwach von ihrem jeweiligen Mittelwert ab. Lediglich der positive Zusammenhang zwischen dem Konsum gemischter Unterhaltung und den *verbalen Kompetenzen* (**0,18**) und der negative Zusammenhang zwischen *niveauarmer Unterhaltung* und der *numerischen Intelligenz* (**-0,11**) liegt mindestens ein Zehntel über bzw. unter dem Durchschnittswert. *Bildungs- und Wettkampfsendungen* stehen in einer leicht positiven Beziehungen zu der *numerischen Intelligenz* (**0,07**), sind aber gleichzeitig minimal negativ mit den *verbalen Kompetenzen* (**-0,09**) verbunden, so dass auch im Fall der weiblichen Probanden davon ausgegangen werden kann, dass diese Sendungen insgesamt in keinem bedeutendem Zusammenhang zu den mathematischen Kompetenzen stehen. Dieses Ergebnis ist insofern schade, als dass ein Großteil der Mädchen (37,7 %) solche Sendungen mit Vorliebe schaut. Bei den Jungen sind die Abweichungen vom männlichen Klassenmittelwert offensichtlicher. Lediglich die Interdependenzen zwischen *niveauarmen Unterhaltungssendungen* und den *verbalen Kompetenzen* liegen im nicht spürbaren Hundertstelbereich. Knapp ein Fünftel der Jungen schaut regelmäßig Fernsehsendungen aus dem Bereich *Wettkampf und Bildung* und im Gegensatz zu den zahlreichen Mädchen ist dies bei den männlichen Jugendlichen scheinbar eine sinnvolle Beschäftigung, die mit klaren positiven Bewertungen sowohl im Hinblick auf die *verbalen* mathematischen *Kompetenzen* (**0,40**) als auch auf die *numerische Intelligenz* (**0,28**) verbunden ist. Betrachtet man nun allerdings die Beziehungen zwischen den beiden anderen Faktoren und den Mathematikleistungen, lassen sich nur noch negative Abweichungen nach unten ablesen. Bereits der präferierte Konsum *gemischter Unterhaltung* ist bei den Jungen mit negativen Auswirkungen auf beide mathematischen Dimensionen verbunden. Da fast zwei Drittel der männlichen Probanden ein solches Fernsehverhalten zeigen, ist dieses Ergebnis durchaus erschreckend. Die berechneten negativen Auswirkungen *niveauarmer Fernsehsendungen* sind dahingegen mit Vorsicht zu genießen. Nur 7 % der Jungen ergötzen sich überwiegend an primitiven Fernsehformaten. Da der Konsum dieser Sendungen vor allem mit negativen Zusammenhängen bzgl. der *numerische Intelligenz* (**-0,24**) einhergeht, ist es als positiv zu erachten, dass weniger als zehn männliche Probanden derartige Sendungen befürworten.

Die Analyse der überdurchschnittlich bewerteten Probanden bestätigt das auf dem arithmetischen Mittelwert basierende Ergebnis für beide Geschlechter. Lediglich *gemischte Unterhaltung* präferierende Mädchen werden zu über der Hälfte positiv beurteilt. Dies gilt vor allem für die *verbalen Kompetenzen* (62,1 %) und spiegelt die positive Abweichung vom weiblichen Klassenmittelwert wieder. Von den zahlreichen *Bildungs- und Wettkampfseherinnen* erhalten nur 47,5 % eine überdurchschnittliche Bewertung in den beiden mathematischen Sektoren. Ebenso werden mehr als 50 % der Mädchen unterdurchschnittlich beurteilt, wenn sie *niveauarmes* Fernsehen konsumieren. Dies gilt für beide Dimensionen, in besonderem Maße allerdings für die *numerische Intelligenz*. Die prozentualen Verteilungen der Jungen bekräftigen vor allem die obigen Aussagen über die positiven Zusammenhänge von *Bildungs- und Wettkampffernsehen* und den Mathematikleistungen. Hier lohnt sich ein genauerer Blick auf die Zusammensetzung der männlichen Probanden.

Insgesamt schneiden 65,4 % bzw. 61,5 % der Jungen mit dieser Präferenz im *verbalen* bzw. *numerischen* Bereich überdurchschnittlich ab. Zwischen 30 und 40 % der Probanden erhalten sogar eine um mehr als einen Bewertungspunkt nach oben abweichende Lehrerbeurteilung. Im umgekehrten negativen Fall gilt dies nur für etwa ein Fünftel der männlichen Jugendlichen. Dementsprechend können die positiven Abweichungen vom männlichen Klassenmittelwert als statistisch korrekt angenommen werden, so dass tatsächlich davon ausgegangen werden kann, dass bei den Jungen der Konsum von *Wettkampfs- und Bildungssendungen* mit besseren mathematischen Leistungen verbunden ist!

Obige Ausführungen können nun anhand von Abbildung 8.10 detailliert nachvollzogen werden kann:

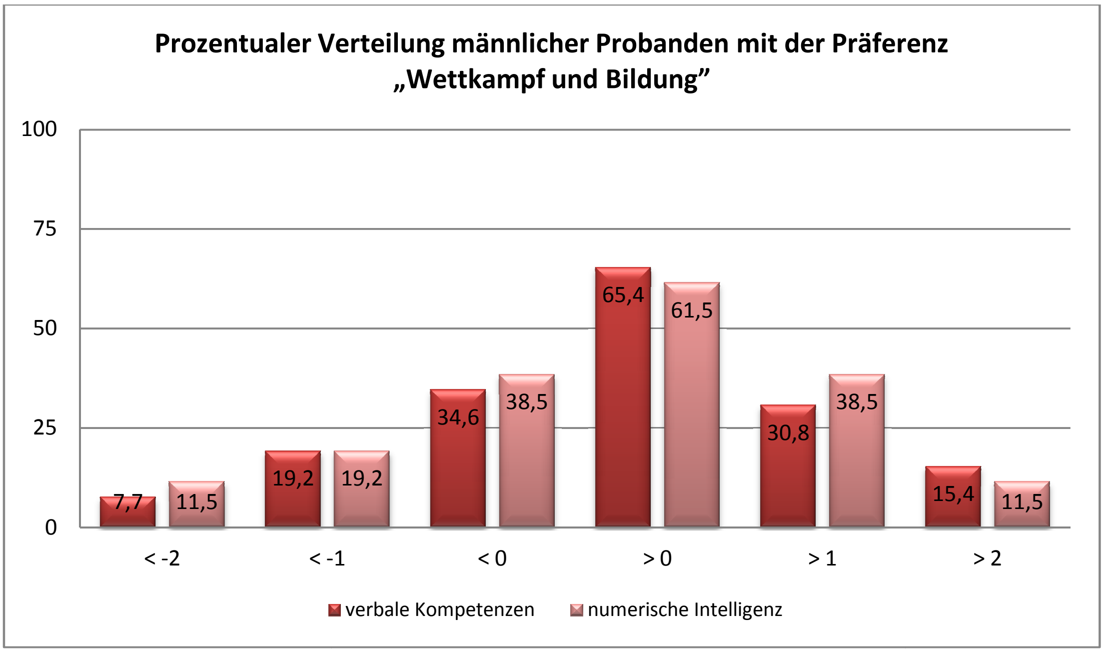


Abbildung 8.10 Prozentuale Verteilung männlicher Probanden mit der Präferenz „Wettkampf und Bildung"

Den abschließenden Block bildet fortan die schulformspezifische Unterscheidung. An dieser Stelle soll erstmals auch ein Blick auf die *humorvolle* und die *musikalische Unterhaltung* geworfen werden. Zwar schauen maximal 2 % der GesamtschülerInnen und nur 1 % der Gymnasiasten derartige Sendungen primär, doch sind diese beiden Faktoren an der Hauptschule von enormer Bedeutung und beispielsweise wesentlich wichtiger als *Bildungssendungen*. Zumindest für die Schülerinnen und Schüler der Hauptschule rechtfertigt sich hier die Miteinbeziehung der beiden letzten Faktoren.

Die schulformabhängige Auswertung beginnt mit einem echten Paukenschlag. Gymnasiasten, die anspruchsvolleres Fernsehen bevorzugen, dies trifft immerhin auf knapp 30 % zu, können daraus kein die mathematischen Fähigkeiten förderndes Wissen ziehen. Ganz im Gegenteil. Für das Gymnasium existiert ein negativer Zusammenhang zwischen dem Faktor *Wettkampf und Bildung* einerseits und den *verbalen* und *numerischen Fertigkeiten* andererseits. Diese negative Beziehung ist sogar mit Abweichungen von bis zu **-0,37** Bewertungspunkten vergleichsweise deutlich ausgeprägt. Über 60 % der Gymnasiasten schauen weniger zielgerichtet fern und bevorzugen die *gemischte Unterhaltung*. Diese Probanden liegen in beiden mathematisch relevanten Dimensionen ein Zehntel über dem klasseninternen Mittelwert. Selbst *anspruchsarme* Gymnasiasten erhalten positiv abweichende Lehrerurteile, wobei dies für weniger als zehn Probanden gilt, so dass die statistische Aussagekraft relativ beschränkt ist. Das wesentliche gymnasiale Ergebnis, welches auch Tabelle 8.4 entnommen werden kann, ist daher eher der negative Zusammenhang zwischen anspruchsvolleren Fernsehsendungen und den entsprechenden Mathematikleistungen.

Tabelle 8.4

Abweichung vom Klassenmittelwert nach Fernsehart und Schulform

	verbale Kompetenz			numerische Intelligenz		
	Gy	Ge	Hs	Gy	Ge	Hs
gemischte Unterhaltung	0,09	0,02	-0,15	0,10	0,09	-0,01
Wettkmapf und Bildung	**-0,37**	0,14	**0,88**	**-0,31**	0,16	**0,74**
niveauarme Unterhaltung	0,21	-0,24	**-0,82**	0,03	-0,42	**-0,89**
humorvolle Unterhaltung	-	-	-0,92	-	-	-1,11
musikalische Unterhaltung	-	-	1,29	-	-	1,36

Auch die meisten Gesamtschülerinnen und -schüler konsumieren vorwiegend Sendungen des ersten Faktors (40 %), wobei der (positive) Zusammenhang gering (*numerisch*) bis minimal (*verbal*) ausfällt. Mehr als ein Drittel der Gesamtschulprobanden favorisiert Fernsehsendungen, die dem Faktor *Wettkampf und Bildung* zuzuordnen sind. Im Gegensatz zu den Gymnasiasten, bei denen das Anschauen dieser Sendungen scheinbar mit keinem Lernzuwachs verbunden ist, stehen die mathematischen Fähigkeiten der Schülerinnen und Schüler von der Gesamtschule zumindest in einer leicht positiven Beziehung. Neben den ersten beiden Faktoren steht an der Gesamtschule auch die *niveauarme* Fernsehkost hoch im Kurs (21 %). Davon sollten die Jugendlichen allerdings besser Abstand nehmen. Der Konsum solcher Sendungen steht in einem negativen Zusammenhang sowohl mit den *verbalen Kompetenzen* (**-0,24**) als auch mit der *numerischen Intelligenz* (**-0,42**).

Besonders interessant ist das Auswertungsergebnis für die Hauptschule, da die Probanden sich hier komplett anders auf die fünf Faktoren verteilen. Zwar belegt die *gemischte Unterhaltung* auch hier die Spitzenposition (31,6 %), doch interessieren sich die Hauptschülerinnen und -schüler nahezu klischeemäßig kaum für *Wettkampfs- und Bildungssendungen* (10,5 %). Stattdessen werden *niveauarme* (21,1 %), *humorvolle* (15,8 %) oder *musikalische Unterhaltungsformate* (21,1 %) vorgezogen. Aus der Perspektive der mathematischen Leistungen ist diese Prioritätensetzung allerdings von Nachteil, da der anspruchsvollere zweite Faktor an der Hauptschule in Zusammenhang mit nach oben abweichenden Leistungen steht, während Probanden, welche *gemischte*, *niveauarme* oder *humorvolle* Unterhaltung präferieren, durchschnittlich bis zu **1,11** Bewertungspunkte unterhalb des Klassenmittelwerts abschneiden. Ein Konsum dieser Sendungen ist mit negativen mathematischen Schulleistungen verbunden!

Schülerinnen und Schüler, die das Fernsehen lediglich zum *Musikhören* nutzen, erhalten dahingegen besonders gute Bewertungen. Hier stellt sich die Frage, inwieweit das Fernsehen in dieser Form noch als visuelles Medium eine Rolle spielt.

Die schulformspezifischen Ergebnisse deuten zusammengefasst darauf hin, dass die Bedeutung des Fernsehprogramms mit nachlassendem Schulniveau zunimmt. Während der Mathematikunterricht des Gymnasiums so anspruchsvoll ist, dass auch *Bildungssendungen* in einem negativen Zusammenhang stehen, erhalten Jugendliche der Gesamt- und vor allem der Hauptschule besonders positive Bewertungen, wenn sie derartige Sendungen konsumieren. Auch die negativen Interdependenzen zu den anspruchsarmen Fernsehsendungen nehmen schulformabhängig zu. Die Hauptschülerinnen und -schüler scheinen nicht in der Lage zu sein, den Aufbau oder den primitiven Inhalt zu durchschauen, und erhalten bei Konsum *niveauarmer Unterhaltung* häufiger unterdurchschnittliche Urteile als dies beispielsweise am Gymnasium der Fall ist, was durch einen Blick auf die prozentuale Verteilung der überdurchschnittlich bewerteten Probanden in Abbildung 8.11 zusätzlich unterstrichen wird:

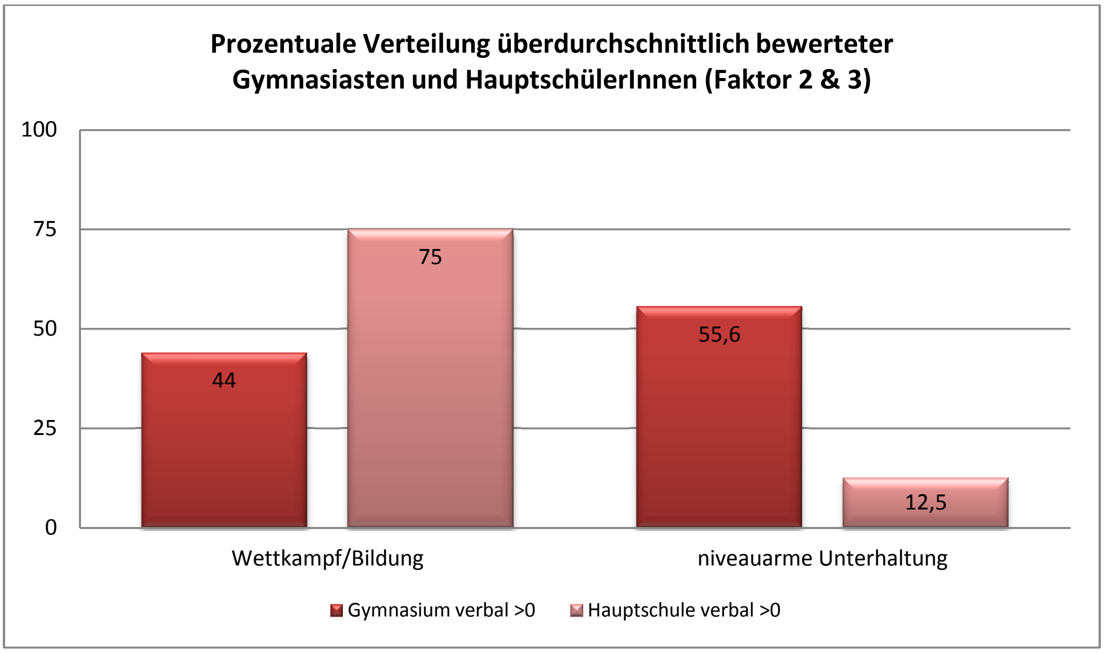


Abbildung 8.11 Prozentuale Verteilung überdurchschnittlich bewerteter Gymnasiasten und HauptschülerInnen (Faktor 2 & 3)

Das Diagramm orientiert sich am (analog übertragbaren) Beispiel der *verbalen Kompetenzen* und vergleicht den Anteil überdurchschnittlich bewerteter Gymnasiasten mit dem Anteil überdurchschnittlich eingestufter Hauptschulprobanden für die beiden Faktoren *Wettkampf und Bildung* und *niveauarme Unterhaltung*. Lediglich 44 % der Jugendlichen, die ein Gymnasium besuchen und häufig *Wettkampfs*- sowie *Bildungsformate* schauen, erhalten positive Bewertungen. An der Hauptschule gilt dies für drei Viertel der Probanden! Dahingegen kann über die Hälfte der *niveauarmes* Fernsehen schauenden Gymnasiasten eine überdurchschnittliche Lehrerbeurteilung vorweisen. Für die Hauptschule ein unmöglich zu erreichender Wert. Hier erhalten gerade einmal 12,5 % dieser Probanden noch eine positive Einschätzung.

Zwischen den *verbalen* sowie *numerischen* Mathematikleistungen und den latenten Faktoren des Fernsehkonsums besteht ein im Vergleich zur täglichen Fernsehdauer nur geringer Zusammenhang. Vor allem die *niveauarme Unterhaltung* ist mit schlechteren Mathematikleistungen verbunden. Positive Interdependenzen von *Wettkampfs- und Bildungssendungen* sind vor allem an der Hauptschule zu beobachten.

8.4 Sonstiges Freizeitverhalten und Mathematikleistung

Nachdem bereits gezeigt worden ist, dass zumindest die isoliert betrachtete Fernsehdauer in einem offensichtlichen Zusammenhang zu den Mathematikleistungen der Probanden steht, während sich die latenten Faktoren des Fernsehkonsums scheinbar nur in einer sehr geringen Beziehung zu dem mathematischen Bildungsstand der Probanden befinden, gilt es abschließend zu untersuchen, ob auch die dritte erhobene Variable, das sonstige fernsehfreie Freizeitverhalten, ähnliche Parallelen aufweist. In dem Abschnitt „Die latenten Faktoren des Freizeitverhaltens" ist es bereits vorab gelungen, die Jugendlichen in die beiden Kategorien *aktiv/kreativ* sowie *passiv/kognitiv anspruchslos* zu unterteilen. Diese beiden latenten Faktoren kennzeichnen die Art des primären Freizeitverhaltens der jeweiligen Probanden und können im Folgenden mit den bekannten Mathematikleistungen in Bezug gesetzt werden.

In einem ersten Schritt betrachte man dazu die in Abbildung 8.12 dargestellten durchschnittlichen Abweichungen der Lehrerbewertungen beider mathematischer Dimensionen vom jeweiligen Klassenmittelwert im Bezug auf das bevorzugte Freizeitverhalten.

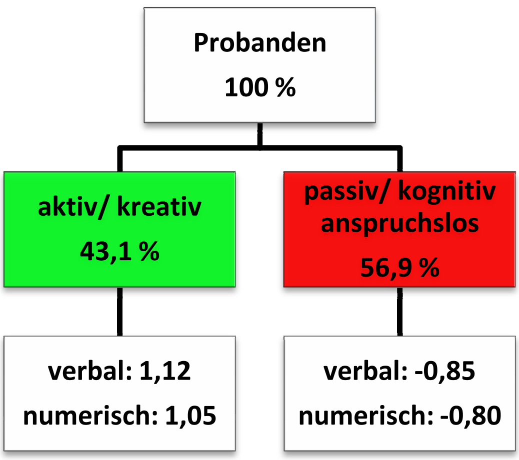


Abbildung 8.12 Freizeitverhalten und Mathematikleistungen

Die Stichprobe besteht zu 43,1 % aus *aktiven* bzw. *kreativen* Jugendlichen. Dementsprechend befürworten die meisten Jugendlichen (56,9 %) ein *passives* und bisweilen *kognitiv anspruchsloses* Freizeitverhalten. Diese Prioritätensetzung bei der individuellen Gestaltung der schulfreien Zeit ist insofern bedauerlich, als dass eine *aktiv/kreative* Freizeitgestaltung mit wesentlich besseren Schulleistungen in Mathematik verbunden ist. Diese Vermutung bestätigt die dritte Ebene des Dendrogramms: Die durchschnittlichen Bewertungen der *verbalen* und *numerischen* mathematischen *Fähigkeiten* der Probanden stehen in einem signifikanten Zusammenhang zu der gewählten primären Freizeitbeschäftigung, getreu dem Sprichwort *nur in einem gesunden Körper wohnt ein gesunder Geist*. Der beobachtete Effekt ist dabei sogar ähnlich prägnant wie bei der Fernsehdauer. Die *aktiv/kreativen* Jugendlichen weichen durchschnittlich um **1,12** (*verbal*) und **1,05** (*numerisch*) Bewertungspunkte nach oben ab, wohingegen die *verbale Kompetenz* der *passiv/kognitiv anspruchslosen* Probanden im Mittel **0,85** und die *numerische Intelligenz* **0,80** Bewertungspunkte unterhalb des jeweiligen Klassendurchschnitts zu finden sind. Da unter anderem die Hobbys „Bücher" und „Schreiben" zu den *aktiven* und *kreativen* Freizeitbeschäftigungen gezählt werden, überrascht es allerdings kaum, dass die *verbalen Kompetenzen* der Probanden in einer leicht stärkeren Beziehung zu dem Freizeitverhalten stehen. Diese Ergebnisse, die Gesamtheit der Probanden betreffend, legen die Vermutung nahe, dass eine *aktive* bzw. *kreative* Freizeitgestaltung durchaus in einem messbaren Zusammenhang zu mathematischen Kompetenzen steht.

Zur grafischen Veranschaulichung der deutlichen Leistungsunterschiede sei auf das Säulendiagramm in Abbildung 8.13 verwiesen.

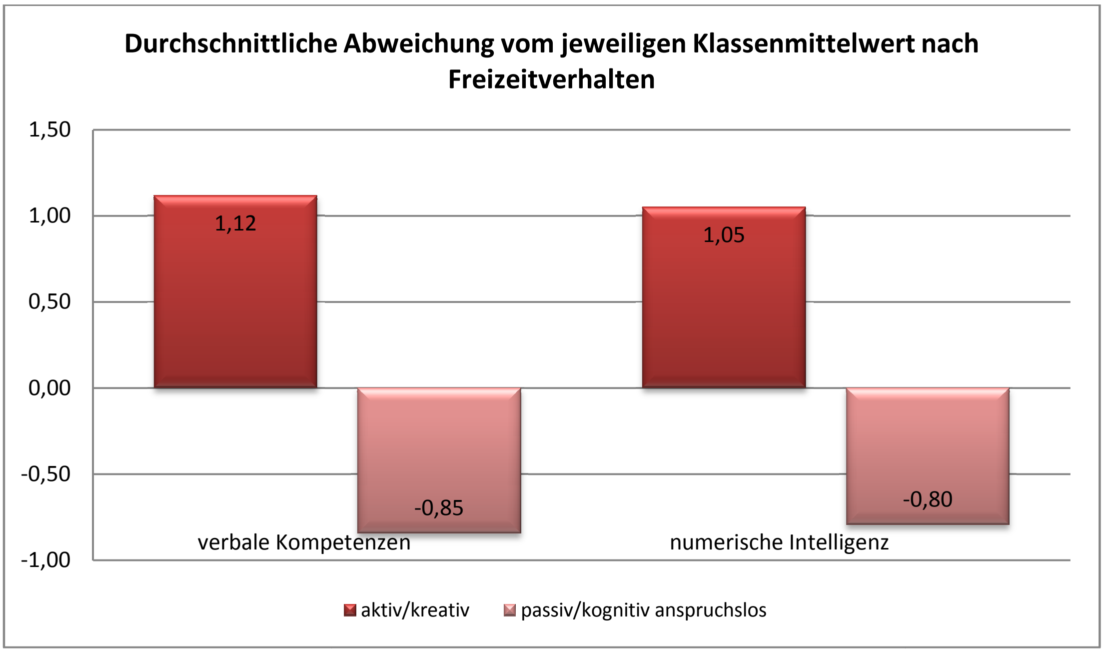


Abbildung 8.13 Durchschnittliche Abweichung vom jeweiligen Klassenmittelwert nach Freizeitverhalten

Nach dieser ersten Tendenz soll, analog zu der bisherigen Ergebnisdarstellung, eine robuste Analyse vorgenommen werden, indem der prozentuale Anteil der überdurchschnittlich bewerteten Probanden in Abhängigkeit vom präferierten Freizeitverhalten bestimmt wird. Das Ergebnis kann in nachfolgender Abbildung 8.14 abgelesen werden und bestätigt abermals den auf dem arithmetischen Mittelwert basierenden Verdacht in vollem Umfang:

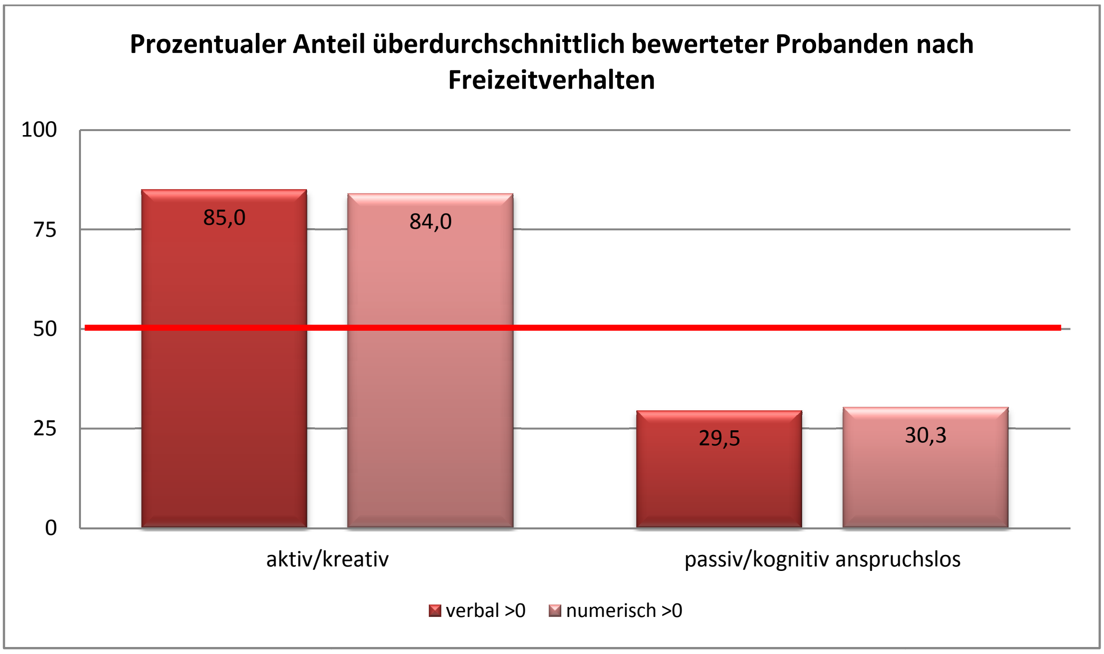


Abbildung 8.14 Prozentualer Anteil überdurchschnittlich bewerteter Probanden nach Freizeitverhalten

Die *aktiv/kreativen* Jugendlichen erhalten demnach nicht nur im Mittel die besten Lehrerurteile, sie werden auch insgesamt zu einem beträchtlichen Anteil überdurchschnittlich gut bewertet. Dies gilt für beide mathematischen Bereiche in (nahezu) gleichem Ausmaß. Etwa 85 % der Probanden mit einer *aktiven* und/oder *kreativen* Freizeitgestaltung werden im Vergleich zum Klassenmittelwert positiv bewertet. Bei den *passiven* sowie *kognitiv anspruchslosen* Untersuchungsteilnehmern gilt dies lediglich für etwa 30 %. Zudem zeigt sich auch in dieser Statistik, dass sich die Beobachtung bei den *verbalen Kompetenzen* leicht verstärkt, wenngleich der Unterschied maximal einen Prozentpunkt beträgt und damit verschwindend gering ist.

Zum ersten Zwischenergebnis des letzten isolierten Auswertungsdurchgangs taugt damit die Erkenntnis, dass ein offensichtlicher Zusammenhang zwischen dem Freizeitverhalten einerseits und der Mathematikleistung der Probanden andererseits besteht. Dieser fällt für die *aktiv/kreativen* Jugendlichen positiv aus, während die *passiv/kognitiv anspruchslosen* Probanden negative Bewertungen erhalten.

Auf bekannte Art und Weise wird dieses Zwischenergebnis um eine geschlechts- und schulformspezifische Dimension ergänzt, wobei mit den Geschlechtern begonnen werden soll.

Tabelle 8.5

Abweichung vom Klassenmittelwert nach Freizeitverhalten und Geschlecht

	verbale Kompetenz		numerische Intelligenz	
	weiblich	männlich	weiblich	männlich
aktiv/kreativ	1,30	0,91	1,20	0,86
passiv/anspruchslos	-0,70	-0,91	-0,64	-0,86

Bereits eine flüchtige Analyse der Tabelle 8.5 verdeutlicht, dass für beide Geschlechter der identifizierte Zusammenhang, inklusive dem größeren Effekt bezüglich der *verbalen Kompetenzen*, bestehen bleibt. Dennoch fallen auch hier zwei geschlechtsspezifische Unterschiede auf. Zum einen werden die mathematischen Fähigkeiten der Mädchen mit einer *aktiv/kreativen* Freizeitbeschäftigung noch besser bewertet als die der Jungen und liegen bis zu **1,30** (*verbal*) Bewertungspunkte oberhalb des weiblichen Klassenmittelwerts. Das Ergebnis lässt sich mit den verschiedenen, zu einem latenten Faktor zusammengefassten Aktivitäten gut erklären. Während die *aktiv/kreativen* Mädchen sich eher intellektuell beschäftigen, indem sie Bücher lesen, Gedichte schreiben oder selbst Musik machen, treiben die Jungen Sport oder spielen Computerspiele. Bedauerlicherweise gehen allerdings nur 35 % der Mädchen überwiegend *aktiv/kreativen* Freizeitbeschäftigungen nach. Das zweite Unterscheidungsmerkmal zwischen den Geschlechtern wird durch die negativen Abweichungen der *passiven* und *kognitiv anspruchslosen* Probanden deutlich. Die Mathematikleistungen dieser Schüler werden in beiden mathematischen Bereichen schwächer beurteilt als die der Schülerinnen. Die negative Abweichung beträgt bei den Jungen durchschnittlich **0,9** Bewertungspunkte, während sie bei den Mädchen im Schnitt maximal **0,70** Punkte unterhalb des Bezugsrahmens liegt.

Die durch den arithmetischen Mittelwert prognostizierten Unterschiede zwischen Jungen und Mädchen werden durch die prozentuale Verteilung überdurchschnittlich bewerteter Probanden nach Freizeitverhalten und Geschlecht in unten stehender Abbildung 8.15 relativiert.

Die Vermutung, dass eine *aktiv/kreative* Freizeitbeschäftigung in einem positiven und eine *passiv/kognitiv anspruchslose* Freizeitgestaltung in einem negativen Zusammenhang zu den Mathematikleistungen beider Geschlechter steht, ist auch Abbildung 8.15 ohne Lupe zu entnehmen und bedarf kaum mehr einer weiteren Bekräftigung. Interessant ist dagegen, dass die im Vergleich zu den Jungen besonders positive Beziehung zwischen den *aktiv/kreativen* Mädchen und den Mathematikleistungen nicht mehr vorhanden ist. Bei den *verbalen Kompeten-*

zen führen die Mädchen nur mit etwa drei Prozentpunkten vor den Jungen, bei der *numerischen Intelligenz* ist der Anteil überdurchschnittlich bewerteter Mädchen sogar geringer als bei den Jungen. Ähnliches gilt für die negative Leistungsdifferenz. Vor allem im *numerischen* Bereich ist ein Unterschied zwischen unterdurchschnittlich bewerteten weiblichen und männlichen Probanden kaum vorhanden.

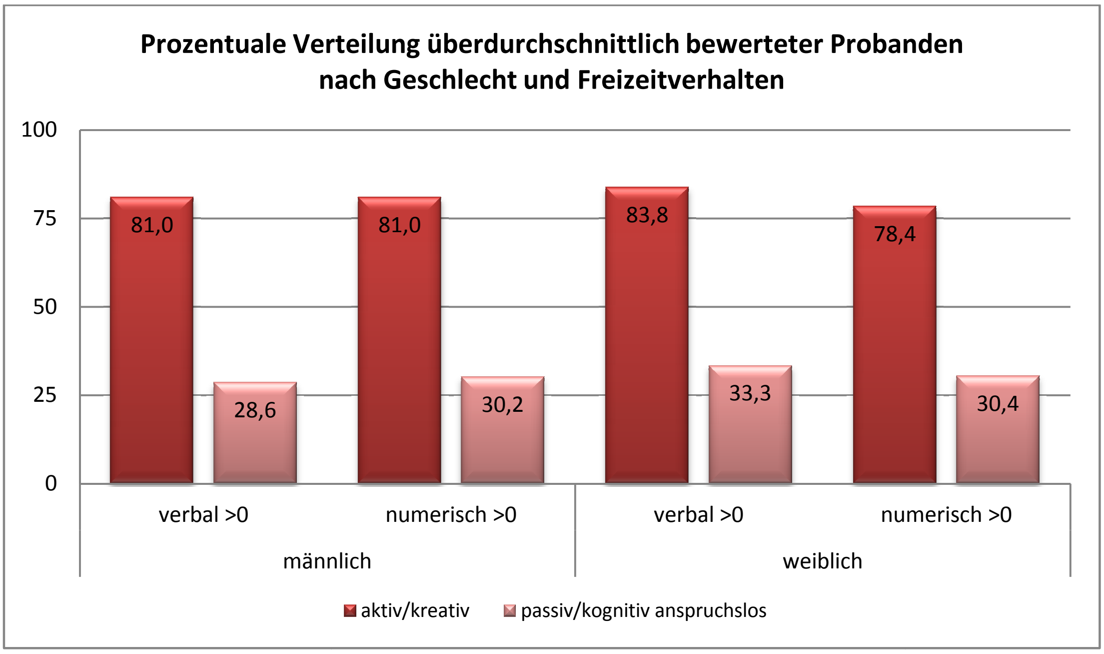


Abbildung 8.15 Prozentuale Verteilung überdurchschnittlich bewerteter Probanden nach Geschlecht und Freizeitverhalten

Dementsprechend sind die identifizierten geschlechtsspezifischen Ungleichheiten mit der mangelnden Robustheit des verwendeten Lagemaßes sowie der genauen Verteilung der Probanden zu erklären, wie Abbildung 8.16 am Beispiel der *numerischen Intelligenz aktiv/kreativer* Probanden offenbart. Zwar erhalten die *aktiv/kreativen* Jungen etwas häufiger eine überdurchschnittliche Bewertung ihrer *numerischen Intelligenz*, doch weicht diese in den wenigsten Fällen um mehr als zwei Punkte nach oben ab. Bei den Mädchen lassen sich wesentlich mehr Ausreißer identifizieren. Beispielsweise haben 32 % eine Lehrerbeurteilung erhalten, die mehr als zwei Bewertungspunkte oberhalb des Vergleichswerts liegt. Selbst Abweichungen von vier (!) Punkten nach oben sind bei den weiblichen Probanden vorhanden. Rechnet man diese Ausreißer heraus, lässt sich die Differenz zwischen Jungen und Mädchen als eher gering einstufen.

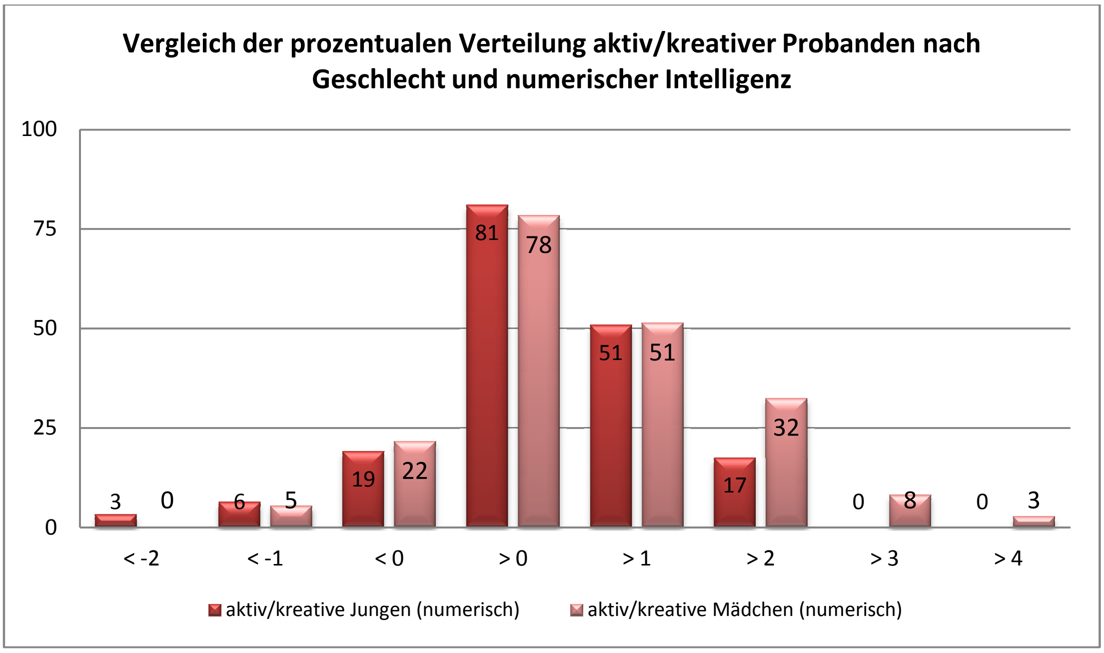


Abbildung 8.16 Vergleich der prozentualen Verteilung aktiv/kreativer Probanden nach Geschlecht und numerischer Intelligenz

Als letztes Unterscheidungsmerkmal dient die von den jeweiligen Probanden besuchte Schulform. Nachdem sowohl für die Fernsehdauer als auch für die latenten Faktoren des Fernsehkonsums bereits auf konkrete schulformbezogene Schwankungen hingewiesen werden konnte, gilt dies offenbar auch für die Interdependenzen zwischen Freizeitverhalten und Mathematikleistung:

Tabelle 8.6

Abweichung vom Klassenmittelwert nach Freizeitverhalten und Schulform

	verbale Kompetenz			numerische Intelligenz		
	Gym	Ge	Hs	Gym	Ge	Hs
aktiv/kreativ	0,88	**1,53**	1,24	0,80	**1,49**	1,18
passiv/anspruchslos	**-1,38**	-0,51	-1,24	**-1,25**	-0,50	-1,18

Zunächst gilt es festzuhalten, dass die Tendenz des positiven Zusammenhangs zwischen *aktiv/kreative* Freizeitbeschäftigung und des negativen Zusammenhangs zwischen *passiv/kognitiv anspruchslosen* Aktivitäten für alle drei Schulformen Gültigkeit besitzt. Die Differenzen liegen im Detail bzw. im Grad der jeweiligen Auswirkung. Der anspruchsvolle Schulalltag am Gymnasium erfordert gewissermaßen auch in der Freizeit ein geistig forderndes Verhalten von den Schülerinnen und Schülern. *Passive* Gymnasiasten weichen in beiden mathematischen Bereichen nämlich besonders deutlich vom Klassenmittelwert nach unten ab.

Die positive Beziehung einer *aktiv/kreativen* Freizeitgestaltung und den Leistungen in Mathematik fällt dagegen vergleichsweise gering aus. Scheinbar ist eine solche Form der Tagesplanung eine notwendige, aber nicht ausreichende Grundvoraussetzung, um überhaupt eine Chance zu haben, gute mathematische Leistungen an einem Gymnasium zu erbringen. Daher verwundert es nicht wirklich, dass nur etwa 40 % der gymnasialen Jugendlichen sich regelmäßig *passiven* und *kognitiv anspruchslosen* Aufgaben widmen, während der Großteil eher am Erlernen eines Musikinstruments o. ä. interessiert ist.

Im Abschnitt „Die latenten Faktoren des Freizeitverhaltens" ist herausgearbeitet worden, dass lediglich ein Viertel der Schülerinnen und Schüler der Gesamtschule ihre Freizeit vorwiegend *aktiv* gestaltet. Selbst an der Hauptschule gilt dies immerhin für die Hälfte der Jugendlichen. Zwar scheinen sich die Gesamtschulprobanden einerseits eine *passive* und *kognitiv anspruchslose* Freizeitgestaltung am ehesten erlauben zu können, da sie durchschnittlich am geringsten vom jeweiligen Klassenmittelwert abweichen, doch führen andererseits die extrem positiven Lehrerurteile der *aktiv/kreativen* Jugendlichen zu der Vermutung, dass in zahlreichen Gesamtschülerinnen und -schülern ein durch mangelndes Freizeitverhalten nicht ausgeschöpftes Potenzial steckt.

Die Jugendlichen von der Hauptschule bilden praktisch das Bindeglied zwischen der Gesamtschule und dem Gymnasium. Die *aktiv/kreativen* Jugendlichen liegen im Durchschnitt zwischen **1,24** (*verbal*) und **1,18** (*numerisch*) Bewertungspunkte oberhalb des Klassenmittelwerts. Bei einem *passiven* und *kognitiv anspruchslosen* Zeitvertreib gelten die gleichen Daten in negativer Richtung.

Der prozentuale Anteil überdurchschnittlich bewerteter Probanden in Abhängigkeit von der Schulform bekräftigt die Ergebnisse in vollem Umfang. Die meisten *aktiv/kreativen* Gesamtschülerinnen und -schüler erfreuen sich tatsächlich einer guten Beurteilung. Dies gilt für 88,5 (*numerisch*) bis 92,3 % (*verbal*), wobei auch verhältnismäßig viele *passiv/anspruchslos* lebende Jugendliche positive Bewertungen erhalten haben. Am Gymnasium greifen die Lehrkräfte bei *passiv/anspruchslosen* Probanden seltener zu einer überdurchschnittlichen Beurteilung, während *aktiv/kreative* Gymnasiasten gute Chancen auf eine positive Beurteilung haben. Gleiches gilt auch für die Jugendlichen der Hauptschule, wie Abbildung 8.17 bekräftigt:

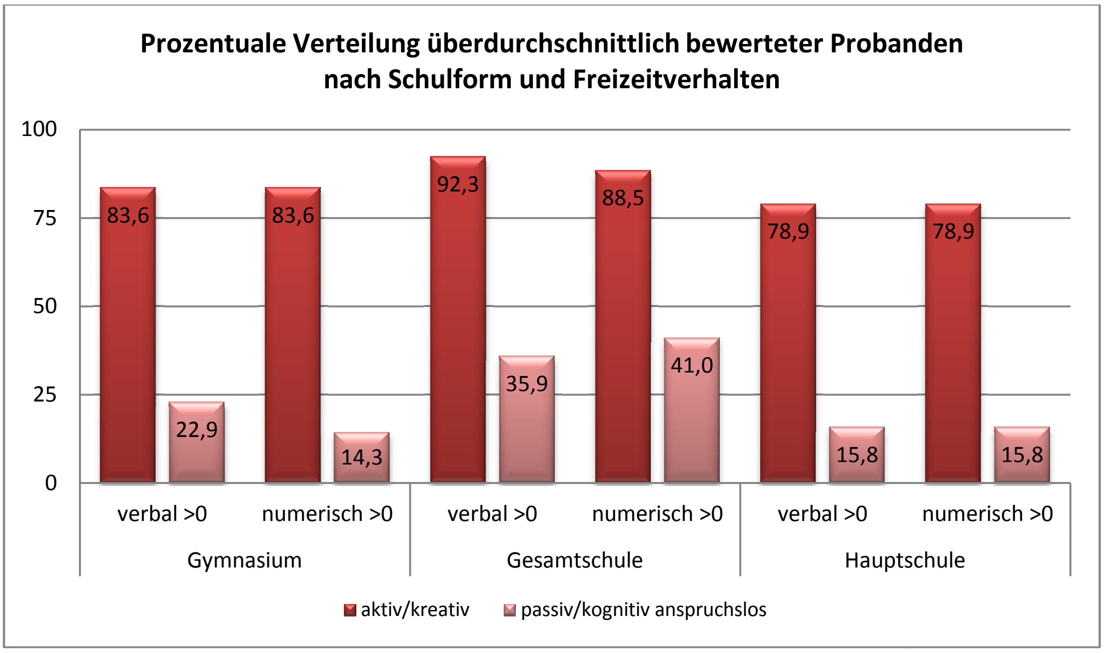


Abbildung 8.17 Prozentuale Verteilung überdurchschnittlich bewerteter Probanden nach Schulform und Freizeitverhalten

Insgesamt bestätigt sich die Vermutung über alle Schulformen und beide Geschlechter hinweg, dass mit einer *aktiv/kreativen* Freizeitgestaltung überdurchschnittliche und mit *passiv/kognitiv anspruchslosen* Freizeitaktivitäten unterdurchschnittliche *verbale* und *numerische* Mathematikleistungen verbunden sind. Es besteht ein offensichtlicher Zusammenhang zwischen Freizeitverhalten und Mathematikleistung!

8.5 Zwischenfazit

Bei den bisherigen Auswertungsschritten sind die drei erhobenen Faktoren isoliert mit den gemessenen Mathematikleistungen der Probanden in Bezug gesetzt worden. Dabei ist jeweils ein eindeutiger Zusammenhang vor allem zwischen der Fernsehdauer sowie zwischen dem Freizeitverhalten und den Mathematikleistungen zu beobachten. Diese Beziehungen gelten sowohl für die *verbalen Kompetenzen* als auch für die *numerische Intelligenz*. Zwischen den konsumierten Sendungen und den mathematischen Kompetenzen scheint dagegen eine wesentlich geringere Verbindung zu bestehen. Dieses Ergebnis unterstützt viele von Manfred Spitzer zitierte Studien und Meinungen, nach denen sich alleine die Fernsehdauer, unabhängig von konsumierten Programminhalten, für schwache Schulleistungen mitverantwortlich zeichnet. Dementsprechend liegen die Fähigkeiten der *Wenig-* und *Mittelseher* im Durchschnitt oberhalb des jeweiligen Klassenmittelwerts, während die *Viel-* und vor allem die *Exzessivseher* schlechter als der entsprechende Vergleichswert bewertet werden. Insbesondere die mathematischen Fähigkeiten der *Wenigseher* werden außerordentlich positiv beurteilt. Die geschlechtsspezifischen Unterschiede fallen dabei gering aus. Bei den Jungen werden vor allem die *verbalen Kompetenzen* bei einer kürzeren Fernsehdauer positiv bewertet, wobei die

Mädchen vergleichsweise weniger schlechte Beurteilungen bei *exzessivem* Fernsehkonsum erhalten. *Wenig* und *mittellang* fernsehende Hauptschülerinnen und -schüler weichen am stärksten nach oben ab. Gymnasiasten mit einem überhöhten Fernsehkonsum erhalten besonders schlechte Bewertungen, so dass davon ausgegangen werden kann, dass die gymnasiale Schullaufbahn auch in der Freizeit geistige Tätigkeiten verlangt.

Der eindeutige Zusammenhang zwischen Freizeitverhalten und Mathematikleistungen besteht für sämtliche Probanden. Insgesamt erhalten *aktiv/kreative* Jugendliche im Durchschnitt Lehrerurteile, die deutlich über dem jeweiligen Klassenmittelwert liegen, während *passiv/kognitiv anspruchslose* Probanden nahezu kollektiv mit negativen Bewertungen rechnen müssen. Die geschlechtsspezifischen Unterschiede sind gering. Der Zusammenhang gilt demnach für beide Geschlechter. Wenn überhaupt, dann erhalten Mädchen bei einer *aktiven* Freizeitgestaltung etwas bessere Urteile und Jungen bei *passiven* Aktivitäten besonders schlechte Lehrerbewertungen. Der Zusammenhang gilt zudem für alle Schulformen, doch schneiden *passive* Gymnasiasten exorbitant schlecht ab. Dieses Ergebnis bekräftigt die oben angestellte Vermutung, ein Gymnasiast müsse auch in der schulfreien Zeit geistig anspruchsvolle Tätigkeiten ausüben, um dem Schulalltag gewachsen zu sein. *Aktive* Gesamt- und Hauptschulprobanden schneiden außerordentlich gut ab.

Zwischen den latenten Faktoren des Fernsehkonsums und den Mathematikleistungen der Probanden besteht allerdings kaum ein spürbarer Zusammenhang. Zwar erhalten Jugendliche, die vorwiegend *niveauarme* Fernsehsendungen präferieren, insgesamt tatsächlich die schlechtesten Lehrerbeurteilungen, doch fällt die mittlere Abweichung nach unten relativ gering aus. Fernsehsendungen des Faktors *Wettkampf und Bildung* stehen in keinem Zusammenhang zu den Mathematikleistungen. Die eine Hälfte der Probanden wird überdurchschnittlich, die andere Hälfte unterdurchschnittlich bewertet. Die Mathematikleistungen der Jungen fallen bei einem anspruchsvollen Fernsehprogramm etwas besser aus. Die Interdependenzen zwischen den bevorzugt geschauten Fernsehsendungen und den Mathematikleistungen der Mädchen sind im Vergleich dazu kaum mehr wahrnehmbar. Die schulformbasierte Auswertung deutet darauf hin, dass der Zusammenhang zwischen den latenten Faktoren des Fernsehkonsums und den Leistungen in Mathematik mit absinkender Schulform ersichtlicher wird. Demnach erhalten Jugendliche der Hauptschule, welche *Wettkampf und Bildung* favorisieren besonders gute Urteile, während sie allerdings bei einer *niveauarmen* Fernsehgestaltung entsprechend schwache Bewertungen zugewiesen bekommen. Die Problematik lässt an der Gesamtschule etwas nach und dreht sich am Gymnasium letztlich um. Eine derartige Beobachtung legt die Interpretation nahe, dass diese Schülerinnen und Schüler von Sendungen wie *Galileo* oder *Wer wird Millionär* wenig Neues und Wichtiges für die Schule lernen. Da Gymnasiasten zusätzlich eher in der Lage sind, *niveauarme* Fernsehserien zu entlarven, stehen ihre Fähigkeiten kaum in einem negativen Zusammenhang zu solchen Formaten.

9. Interdependenzen zwischen Fernsehverhalten und Mathematikleistung: die kombinierte Auswertung

9.1 Die Methoden zur endgültigen Ermittlung der Hierarchie

Nach den isolierten Auswertungsetappen gilt es im Folgenden zu untersuchen, ob sich die Zusammenhänge bestätigen, wenn die drei Variablen hierarchisch kombiniert werden. Erst dann ist es möglich, detaillierte und prägnante Schlussfolgerungen zu ziehen. Der endgültigen Vorgehensweise zur Ermittlung dieser Hierarchie liegen dabei zwei verschiedene Methoden zugrunde.

Auf Grundlage der durchgeführten Befragungen lässt sich einerseits in kanonischer, rein deskriptiver Art und Weise eine Hierarchie der erfragten Daten ermitteln, indem entsprechend der Reihenfolge der vorgelegten Fragen, die befragten Schülerinnen und Schüler divisiv in immer feiner werdende Cluster aufgeteilt werden. Bei einem vorgegebenen Distanzmaß „d", das lediglich ordinale Signifikanz hat, gibt es auf der anderen Seite zwei „optimale" Verfahren der hierarchischen Clusteranalyse, welche zum einen die „d" eindeutig bestimmte maximal subdominierende Ultrametrik liefern und zum anderen alle „d" minimal dominierenden Ultrametriken. Dies sind das Single-Linkage und das Complete-Linkage Verfahren. Beide Verfahren sind agglomerativ. Die zugehörigen Fusionskriterien sind durch

$$D\left(A,B\right) := \min_{\left(a,b\right)\in A\times B} d\left(a,b\right) \quad \left(\textit{Single-Linkage}\right)$$

bzw.

$$D\left(A,B\right) := \max_{\left(a,b\right)\in A\times B} d\left(a,b\right) \quad \left(\textit{Complete-Linkage}\right)$$

definiert. Mögliche Distanzmaße, auf denen diese Verfahren, bei zugrunde liegenden reellen Daten, basieren könnten, sind der Euklidische Abstand $\left\|\cdot\right\|$ bzw. die City-Block-Metrik $\left\|\cdot\right\|_c$.

Beide Distanzmaße führen in der Regel bei Benutzung des Single-Linkage bzw. Complete-Linkage Verfahrens zu völlig unterschiedlichen Hierarchien. Werden diese Distanzen für eine fest gewählte natürliche Zahl $n \geq 2$ für Tupel $x \in \left(\mathbb{R}^{\geq 0}\right)^n$ und $y \in \left(\mathbb{R}^{\geq 0}\right)^n$ gemessen, dann erfüllen sie darüber hinaus nicht für jeden Ordnungsautomorphismus T auf $\left(\mathbb{R}^{\geq 0}, \leq\right)$ die Äquivalenz $\|x\| \leq \|y\| \Leftrightarrow \|\mathrm{T}(x)\| \leq \|\mathrm{T}(y)\|$ bzw. $\|x\|_c \leq \|y\|_c \Leftrightarrow \|\mathrm{T}(x)\|_c \leq \|\mathrm{T}(y)\|_c$. Dies meint, dass weder der Euklidische Abstand noch die City-Block-Metrik i.a. ordinale Signifikanz haben. Daher soll im folgenden Satz unter der Voraussetzung, dass die Ausgangsdaten nicht auf Intervallskalenniveau gemessen wurden, zum einen geklärt werden, für welche Teilmengen $\mathbf{S}$ des $\left(\mathbb{R}^{\geq 0}\right)^n$ der Euklidische Abstand einerseits und die City-Block-Metrik andererseits in Bezug auf das Single-Linkage bzw. Complete-Linkage Verfahren zu gleichen Hierarchien führen und zum anderen für welche Teilmengen $\mathbf{S}$ des $\left(\mathbb{R}^{\geq 0}\right)^n$ der Euklidische Abstand einerseits und die City-Block-Metrik andererseits ordinale Signifikanz haben.

Zur Formulierung des entsprechenden Satzes benötigen wir folgende ad hoc Notationen, die insbesondere das mögliche ordinale Niveau der Ausgangsdaten reflektieren:

Eine Untergruppe $(\mathbb{T}, \circ)$ der Gruppe $\mathbf{Aut}\left(\mathbb{R}^{\geq 0}, \leq\right)$ aller Ordnungsautomorphismen T auf $\left(\mathbb{R}^{\geq 0}, \leq\right)$ heißt *lokal nicht-linear*, wenn sie für alle positiven reellen Zahlen $a < c < d < b$ und $a < g < h < b$ oder alle positiven reellen Zahlen $c < a < b < d$ und $g < a < b < h$ die folgende entsprechende Bedingung erfüllt.

LN1: Es existiert ein Ordnungsautomorphismus $\mathrm{T} \in \mathbb{T}$, für den $\mathrm{T}(w) = w$ für alle nichtnegativen reellen Zahlen $w \leq a$ und $\mathrm{T}(z) = z$ für alle positiven reellen Zahlen $b \leq z$ sowie $\mathrm{T}(c) = g$ und $\mathrm{T}(d) = h$ gilt

oder

LN2: Es existiert ein Ordnungsautomorphismus $\mathrm{T} \in \mathbb{T}$, für den $\mathrm{T}(v) = v$ für alle reellen Zahlen $a \leq v \leq b$ sowie $\mathrm{T}(c) = g$ und $\mathrm{T}(d) = h$ gilt.

Für alle natürlichen Zahlen $n \geq 2$ heißt eine Teilmenge $\mathbf{S}$ des $\left(\mathbb{R}^{\geq 0}\right)^n$ dann $\mathbb{T}$-*abgeschlossen*, wenn mit jedem Tupel $x = \left(x_1, \ldots, x_n\right) \in \mathbf{S}$ und jedem Ordnungsautomorphismus $\mathrm{T} \in \mathbb{T}$ auch das Tupel $\mathrm{T}(x) := \left(\mathrm{T}\left(x_1\right), \ldots, \mathrm{T}\left(x_n\right)\right)$ in $\mathbf{S}$ enthalten ist.

Darüber hinaus sei für jedes Tupel $x=\left(x_1,\dots,x_n\right)\in\left(\mathbb{R}^{\geq 0}\right)^n$ die natürliche Zahl $N\left(x\right)$ durch $N\left(x\right):=\left|\left\{1\leq k\leq n\,\middle|\,x_k\neq 0\right\}\right|$ definiert. Für jede positive reelle Zahl a und jede natürliche Zahl $0\leq m\leq n$ setzen wir dann $\left\{0,a\right\}_m^n:=\left\{x\in\left\{0,a\right\}^n\,\middle|\,N\left(x\right)=m\right\}$.

Mit Hilfe dieser Notationen lässt sich dann folgender Satz beweisen, wobei der entsprechende Beweis im Anhang einzusehen ist.

Satz 9.1:
Seien $n\geq 2$ eine natürliche Zahl, $\left(\mathbb{T},\circ\right)$ eine lokal nicht-lineare Untergruppe von $\mathbf{Aut}\left(\mathbb{R}^{\geq 0},\leq\right)$ und $\mathbf{S}$ eine $\mathbb{T}$-abgeschlossene Teilmenge des $\left(\mathbb{R}^{\geq 0}\right)^n$. Dann sind folgende Aussagen äquivalent:

$\left(\mathrm{i}\right)$ *Für alle Tupel $y\in\mathbf{S}$ und $z\in\mathbf{S}$ gilt die Äquivalenz $\left\|y\right\|\leq\left\|z\right\|\Leftrightarrow\left\|y\right\|_c\leq\left\|z\right\|_c$.*

$\left(\mathrm{ii}\right)$ *Für alle Tupel $y\in\mathbf{S}$ und $z\in\mathbf{S}$ gilt die Implikation $\left\|y\right\|\leq\left\|z\right\|\Rightarrow\left\|y\right\|_c\leq\left\|z\right\|_c$.*

$\left(\mathrm{iii}\right)$ *Für alle Tupel $y\in\mathbf{S}$ und $z\in\mathbf{S}$ gilt die Implikation $\left\|y\right\|_c\leq\left\|z\right\|_c\Rightarrow\left\|y\right\|\leq\left\|z\right\|$.*

$\left(\mathrm{iv}\right)$ *Es existiert eine feste natürliche Zahl $1\leq m\leq n$ mit $\mathbf{S}\subset\bigcup\limits_{a\in\mathbb{R}^{\geq 0}}\left\{0,a\right\}_m^n$.*

Darüber hinaus haben sowohl $\left\|\cdot\right\|$ als auch $\left\|\cdot\right\|_c$ auf $\mathbf{S}$ ordinale Signifikanz.

Satz 9.1 lässt sich durch die Aussage vertiefen, dass sowohl $\left\|\cdot\right\|$ als auch $\left\|\cdot\right\|_c$ genau dann ordinale Signifikanz haben, wenn eine feste natürliche Zahl $1\leq m\leq n$ existiert, so dass $\mathbf{S}$ in $\bigcup\limits_{a\in\mathbb{R}^{\geq 0}}\left\{0,a\right\}_m^n$ enthalten ist. Der Beweis für diese Äquivalenz wird hier nicht mehr geführt. Satz 9.1 ist einschränkend genug. Andererseits gilt jedoch, dass für jede positive reelle Zahl a auf $\left\{0,a\right\}^n$ die in Aussage $\left(\mathrm{i}\right)$ von Satz 9.1 behauptete Äquivalenz gilt. Also legt es Satz 9.1 nahe, dem Geist der Ausführungen zur Faktorenanalyse zu folgen und dichotomisierte Variable zu betrachten. Daher werden im Folgenden zur Ergänzung des eingangs erwähnten divisiven deskriptiven Verfahrens sowohl das Single-Linkage als auch das Complete-Linkage Verfahren für sich im Rahmen der Befragungen kanonisch ergeben habende 0-1-Daten angewendet.

9.2 Die Gruppen des Fernseh- und Freizeitverhaltens

Nachdem in Abschnitt 9.1 der methodische Hintergrund zur endgültigen Ermittlung der Hierarchie beschrieben wurde, sollen nun die Gruppen des Fernseh- und Freizeitverhaltens auf die zwei beschriebenen Arten gebildet werden. Zunächst beginne man mit dem deskriptiven Vorgehen.

Dabei wird deutlich, dass insgesamt $4 \cdot 5 \cdot 2 = 40$ verschiedene Fernseh- und Freizeitoptionen bestehen, da vier Gruppen der Fernsehdauer mit fünf latenten Faktoren des Fernsehkonsums und zwei möglichen Freizeittypen verknüpft werden. Dabei kann es durchaus vorkommen, dass einige der 40 Fälle durch besonders viele und andere durch besonders wenig Jugendliche repräsentiert werden. Um die statistische Aussagekraft der Studie nicht infrage zu stellen, sollen daher lediglich diejenigen Kombinationen Beachtung finden, die von mindestens acht Probanden bzw. 3,4 % der Stichprobe repräsentiert werden, da das Risiko einer Fehlinterpretation der erhobenen Daten bei absinkender Gruppengröße zunimmt. Des Weiteren ist es nur schwer möglich 40 verschiedene Eventualitäten miteinander zu vergleichen, ohne die Übersicht zu verlieren und in einem nicht enden wollenden Wirrwarr aus Daten verloren zu gehen.

Diese Vorgehensweise führt zu dem in Abbildung 9.1 dargestelltem Dendrogramm, welches die finalen Kombinationen der drei Variablen darstellt und gleichzeitig direkt Bezug auf die zu vergleichenden Mathematikleistungen der Probanden nimmt:

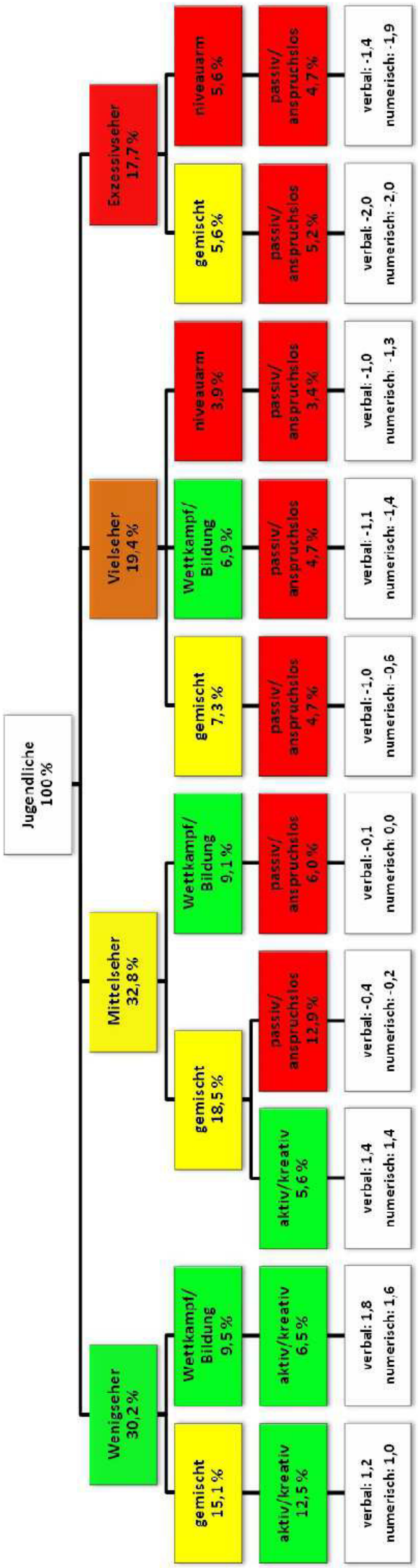


Abbildung 9.1 Fernsehdauer, latente Faktoren des Fernsehkonsums, Freizeitverhalten und Mathematikleistungen

Ausgehend von der vollständigen Stichprobe, bestehend aus 100 % der Probanden, teilt sich der Pfad in einer ersten Etappe in vier, die verschiedenen Gruppen der Fernsehdauer repräsentierenden Wege auf. Sämtliche Fernsehdauertypen beinhalten demnach Fälle, die aus mindestens acht Probanden bestehen.

Auf der nächsten Stufe werden die in Abhängigkeit von der Fernsehdauer geschauten latenten Faktoren des Fernsehkonsums dem Dendrogramm hinzugefügt. An dieser Stelle werden nicht repräsentative Faktoren erstmals unterdrückt. Die meisten *Wenigseher* mögen Sendungen aus den Kategorien *gemischte Unterhaltung* (15,1 %) und *Wettkampf und Bildung* (9,5 %). Diese Probanden schauen dementsprechend nicht nur zeitlich gemäßigt fern, sondern entscheiden sich zu einem Großteil auch für die augenscheinlich niveauvolleren Fernsehsendungen. Ähnliches gilt für die *Mittelseher*, welche ebenfalls eher die *gemischte Unterhaltung* (18,5 %) und *Wettkampf und Bildung* (9,1 %) bevorzugen. Im Vergleich zu den Wenigsehern verzeichnet die *gemischte Unterhaltung* bei den *Mittelsehern* ein Plus um 3,4 Prozentpunkte, während *Wettkampf- und Bildungssendungen* ein leichtes Minus von 0,4 Prozentpunkten hinnehmen müssen. In beiden Gruppen spielt die *niveauarme Unterhaltung* kaum eine Rolle. Nur etwa 10 % sämtlicher *Wenig*- und *Mittelseher* präferieren Doku-Soaps und Castingshows. Dies ändert sich allerdings bei den *Vielsehern*. 3,9 % der Stichprobe schauen folgerichtig nicht nur lange fernsehen, sondern favorisieren gleichzeitig besonders *niveauarme* Fernsehkost. Nichtsdestotrotz bleiben die *gemischte Unterhaltung* (7,3 %) und *Wettkampf und Bildung* (6,9 %) auch unter den *Vielsehern* beliebt und dominant. Erst bei den *exzessiv* fernsehenden Jugendlichen wendet sich das Blatt vollkommen. Niveauvolle Sendungen werden vergebens gesucht. *Niveauarme* und *gemischte Unterhaltung* (je 5,6 %) bilden das zentrale Fundament dieser Probanden. Die Tendenz deutet folgerichtig darauf hin, dass mit zunehmender Fernsehdauer auch die Qualität der konsumierten Programminhalte nachlässt.

Nachdem während der zweiten Etappe die Gruppen der Fernsehdauer mit den latenten Faktoren des Fernsehkonsums zusammengeführt werden konnten, gilt es abschließend das Freizeitverhalten der Probanden den bisher entstandenen Pfadwegen hinzuzufügen. Anhand dessen wird deutlich, dass eine geringe Fernsehdauer in vielen Fällen nicht nur mit einem anspruchsvolleren Fernsehprogramm, sondern auch mit einer *aktiv/kreativen* Freizeitgestaltung zusammenhängt, während eine hohe tägliche Fernsehdosis neben einer niveauärmeren Sendungsauswahl zusätzlich häufig *passive und kognitiv anspruchslose* Hobbys mit sich führt. Beide identifizierten *Wenigsehergruppen* beschäftigen sich demnach in ihrer Freizeit vorwiegend *aktiv*. Bei den *Mittelsehern* lässt dieser Anteil bereits nach. Während sich *Bildungs- und Wettkampfsendungen* favorisierende Jugendliche bereits zu zwei Dritteln *passiv und kognitiv anspruchslos* verhalten, wählen Probanden, die den Faktor der *gemischten Unterhaltung* vorziehen sogar in 70 % der Fälle solche Freizeitbeschäftigungen aus. Für Schülerinnen und Schüler, die *viel* Zeit vor dem Fernseher verbringen oder sogar *exzessiv* fernsehen, sind *ak-*

tiv/kreative Freizeitaktivitäten schließlich kaum mehr von Bedeutung. Für alle fünf Gruppen gilt unabhängig von den latenten Faktoren des Fernsehkonsums, dass sich zwischen 65 und 93 % der Probanden in der schulfreien Zeit *passiven und kognitiv anspruchslosen* Beschäftigungen widmen.

Abschließend werden die von den Lehrerinnen und Lehrern beurteilten Mathematikleistungen auf der letzten Stufe dem Dendrogramm hinzugefügt. In den beiden mathematischen Bereichen *numerische Intelligenz* und *verbale Kompetenzen* handelt es sich dabei jeweils um das arithmetische Mittel des Lehrerurteils.

Insgesamt können die Probanden auf 34 verschiedene Gruppen aufgeteilt werden. Aufgrund der Voraussetzung, nur diejenigen Gruppen zu betrachten, welche mindestens 8 Probanden beinhalten, enthält das vervollständigte Dendrogramm zehn unterschiedliche Pfade!

> *Noch erwähnt werden soll abermals, dass der hierarchische Aufbau des Dendrogramms lediglich von der Reihenfolge der erhobenen Daten abhängt. Es ist nicht angezeigt, auf Grundlage der Hierarchie Interpretationen über die Wichtigkeit der einzelnen Variablen anzustellen. Das Dendrogramm dient lediglich dazu, die verschiedenen Gruppen übersichtlich und prägnant darzustellen. Bei einer anderen Reihenfolge der Variablen würden identische Gruppen entstehen! Auch die mathematischen Leistungen der Probanden am Ende des Dendrogramms stellen abermals lediglich Zusammenhänge dar. Eine Einflussrichtung darf nicht aus dem Aufbau der grafischen Darstellung abgeleitet werden.*

Zur Bestätigung der Methodenunabhängigkeit der durch das Dendrogramm 9.1 deskriptiv ermittelten Gruppen, werden die Probanden im Folgenden auf Basis der zwei Clustertechniken Single-Linkage und Complete-Linkage entsprechend Satz 9.1 ebenfalls zu Gruppen mit einem ähnlichen Fernseh- und Freizeitverhalten zusammengefasst.

Um zunächst den Forschungsansatz dieser Arbeit beizubehalten, werden die erhobenen Daten für die drei Variablen in 0-1-Daten übertragen. Die Anzahl der Komponenten in den 0-1-Tupeln stimmt mit der Anzahl der verschiedenen Gruppen, die zu einem Fragenkomplex ermittelt werden konnten, überein. Da beispielsweise vier verschiedene Gruppen der täglichen Fernsehdauer identifiziert wurden, bestehen die Tupel, welche die Fernsehdauer beschreiben, logischerweise aus vier Komponenten. Tabelle 9.1 vermittelt einen Überblick über sämtliche in 0-1-Daten transformierte Gruppen.

Tabelle 9.1

In 0-1-Daten transformierte Gruppen

Fernsehdauer		Fernsehart		Freizeitverhalten	
Wenigseher	(1,1,1,1)	gemischt	(1,1,1,1,1)	aktiv/kreativ	(1,1)
Mittelseher	(0,1,1,1)	Wettkampf/Bildung	(0,1,1,1,1)	passiv	(0,1)
Vielseher	(0,0,1,1)	niveauarm	(0,0,1,1,1)		
Exzessivseher	(0,0,0,1)	musikalisch	(0,0,0,1,1)		
		humorvoll	(0,0,0,0,1)		

Nun kann das Antwortverhalten eines jeden Probanden in einem elf-elementigen Tupel $P_X = (\underbrace{x_1, x_2, x_3, x_4}_{\text{Fernsehdauer}}, \underbrace{x_5, x_6, x_7, x_8, x_9}_{\text{Fernsehart}}, \underbrace{x_{10}, x_{11}}_{\text{Freizeit}})$ mit $x_4 = x_9 = x_{11} = 1$ und $x_1, x_2, x_3, x_5, x_6, x_7, x_8, x_{10} \in \{0,1\}$, wobei $x_1 \leq x_2 \leq x_3 \wedge x_5 \leq \ldots \leq x_8$ gelten, zusammengefasst werden. Dabei beschreiben die ersten vier Komponenten die präferierte Fernsehdauer, die Komponenten fünf bis neun geben Auskunft über die bevorzugte Fernsehart und die letzten beiden Komponenten artikulieren das Freizeitverhalten. Ein Proband, welcher *mittellang niveauarme Unterhaltung* sieht und sich in der übrigen Freizeit *passiven* Aktivitäten widmet (wie Proband 29), kann dann zum Beispiel durch folgendes Tupel dargestellt werden: $P_{029} = (\underbrace{0,1,1,1}_{\text{Mittelseher}}, \underbrace{0,0,1,1,1}_{\substack{\text{niveauarme} \\ \text{Unterhaltung}}}, \underbrace{0,1}_{\text{passiv}})$.

Im Anschluss an die Transformation der erhobenen Daten sämtlicher Probanden in die entsprechenden 0-1-Daten können die beiden angekündigten clusteranalytischen Verfahren auf den Datensatz angewendet werden. Gemäß Satz 9.1 ist das Ergebnis der Clusteranalysen nun unabhängig von dem gewählten Distanzmaß und kann an nachfolgender Abbildung 9.2 abgelesen werden. Fasst man nun diejenigen Probanden zu je einem gemeinsamen Cluster zusammen, die zueinander den kleinst möglichen Abstand „0" aufweisen, in Abbildung 9.2 handelt es sich dabei um die gebildeten Gruppen unterhalb der roten Trennlinie, liefern sowohl das Single-Linkage als auch das Complete-Linkage Verfahren exakt die selben Cluster wie das deskriptive Vorgehen zuvor, so dass logischerweise die Anzahl aller Cluster (34) sowie die statistisch relevanten Gruppen (in Abbildung 9.2 grün markiert) ebenfalls mit den Ergebnissen der divisiv erstellten Hierarchie übereinstimmen. Die Zusammensetzung der Gruppen ist dementsprechend unabhängig von der gewählten Methode.

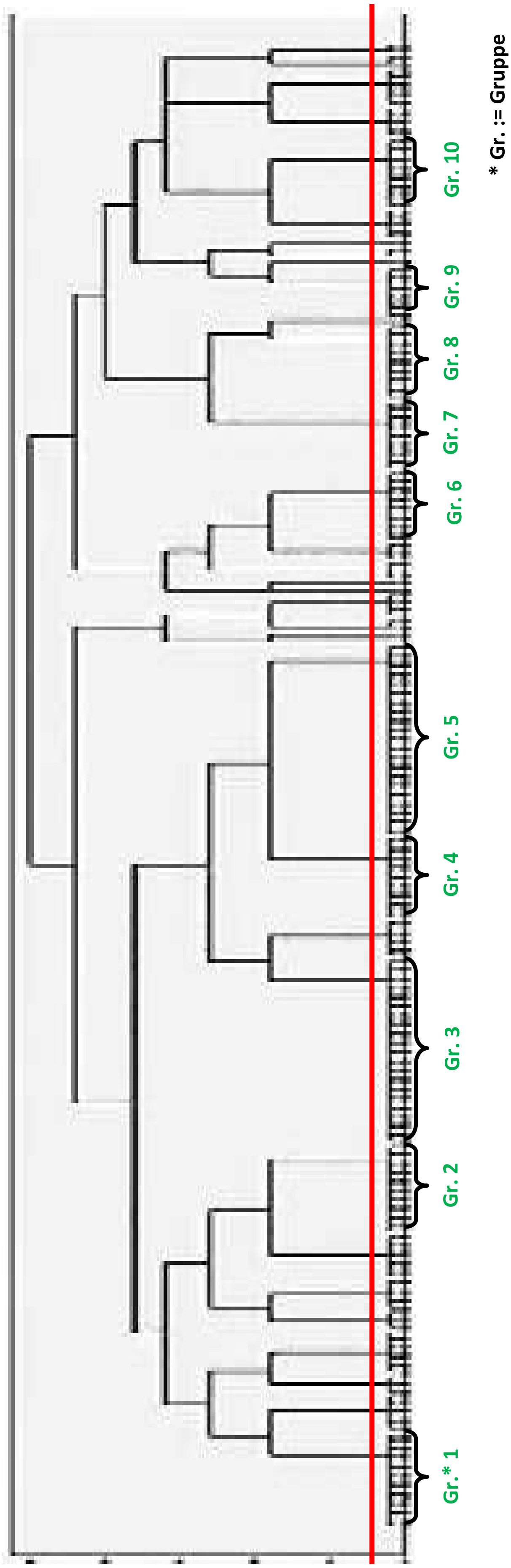


Abbildung 9.2 Ergebnis der Clusteranalyse

Eine Übereinstimmung der durch das Complete-Linkage Verfahren gewonnenen Cluster der Hierarchie mit der durch das Single-Linkage Verfahren gewonnenen Cluster der Hierarchie ist äquivalent zu der Aussage, dass das Ausgangsdistanzmaß „d" bereits eine Ultrametrik war. Dieser Sachverhalt lässt sich mathematisch wie folgt präzisieren.

Bezeichnet man zwei Distanzmaße d und d' als *monoton-äquivalent*, wenn ein Ordnungsautomorphismus $\varphi:(\mathbb{R},\leq)\mapsto(\mathbb{R},\leq)$ existiert, so dass $d'=\varphi\circ d$ gilt und kürzt man darüber hinaus die zu d gehörige Äquivalenzklasse wie üblich mit d ab, so gilt folgender elementarer Satz, dessen trivialer Beweis der Kürze halber hier ausgelassen wird.

Satz 9.2:

Für alle Distanzmaße $d^n \in d$ sind die durch das Complete-Linkage Verfahren und die durch das Single-Linkage Verfahren gewonnenen Distanzmaße genau dann äquivalent, wenn d aus Ultrametriken besteht.

Bevor die *verbalen* und *numerischen* Mathematikleistungen der Probanden und die ermittelten Pfade bzw. Cluster nun in Bezug gesetzt werden, sollen vorab einige interessante Auffälligkeiten zusammengefasst werden. Wie bereits beim deskriptiven Vorgehen erläutert, werden dabei lediglich die zehn statistisch und inhaltlich relevanten Gruppen, welche mehr als 3,4 % der Probanden beinhalten, betrachtet. Die weitere Auswertung bezieht sich aufgrund der hervorzuhebenden Übersichtlichkeit auf die deskriptiv entwickelte Hierarchie.

Die drei Faktoren (Fernsehdauer, Fernsehart, Freizeitverhalten) bedingen sich gegenseitig. Ein zeitlich geringer Fernsehkonsum steht häufig in Verbindung mit einer qualitativ ansprechenden Sendungsauswahl und einer auch ansonsten *kreativen* und *aktiven* Freizeitbeschäftigung (Abbildung 9.1: Pfad 1, 2 und 3). Mit zunehmender Fernsehdauer nimmt sowohl das Niveau der bevorzugten Programminhalte als auch die Bedeutung *aktiv/kreativer* Hobbys kontinuierlich ab (Abbildung 9.1: ab Pfad 4). Ab einer *langen* täglichen Fernsehdosis stehen demnach vor allem *passive* und *kognitiv anspruchslose* Freizeitbeschäftigungen und Fernsehserien der Faktoren *gemischte Unterhaltung* und *niveauarme Unterhaltung* im Mittelpunkt der schulfreien Zeit (Abbildung 9.1: Pfad 9 und 10). Diese Ergebnisse führen zu der Vermutung, dass diejenigen Jugendlichen, die zu den ersten drei Pfaden gehören, die besseren Mathematikleistungen aufweisen sollten, als die Schülerinnen und Schüler ab Pfad vier. Diese Vermutung gilt es im Folgenden zu untersuchen.

9.3 Fernsehdauer, latente Faktoren des Fernsehkonsums, Freizeitverhalten und Mathematikleistung

Zur Überprüfung der obigen Vermutung werden im Folgenden die am jeweiligen Klassenmittelwert z-standardisierten Lehrerbewertungen sowohl für die *verbalen Kompetenzen* als auch für die *numerische Intelligenz* mit den zehn inhaltlich und statistisch relevanten Pfaden, die in Abschnitt 9.2 ermittelt wurden, in Bezug gesetzt. Die positiven und negativen Abweichungen, die bereits am Ende des jeweiligen Pfades in Abbildung 9.1 festgehalten wurden, werden im Sinne einer besseren Übersichtlichkeit nun in das nachfolgende Säulendiagramm, welches in Abbildung 9.3 auf der nächsten Seite dargestellt ist, übertragen und absteigend angeordnet.

Dem Diagramm kann entnommen werden, dass drei der zehn Pfade in einem positiven Zusammenhang zu den beiden Dimensionen der Mathematikleistungen stehen. Es handelt sich wenig überraschend um die beiden *Wenigsehergruppen* und die einzige *Mittelsehergruppe*, deren Probanden zusätzlich einer *aktiv/kreativen* Freizeitbeschäftigung nachgehen. Dabei erzielen die im Vergleich zum jeweiligen Klassenmittelwert besten Lehrerurteile tatsächlich diejenigen Probanden, die *wenig* fernsehen, sich dabei auf *Wettkampfs- und Bildungssendungen* beschränken und in der Freizeit *aktiv/kreativen* Tätigkeiten nachgehen. Die *verbalen Kompetenzen* weichen durchschnittlich mit **1,8** und die *numerische Intelligenz* im Schnitt mit **1,4** Bewertungspunkten deutlich nach oben ab. Für ein wenig Verwunderung sorgt möglicherweise die Erkenntnis, dass die *Mittelsehergruppe* sich vor der zweiten *Wenigsehergruppe* einsortiert. Diese beiden Pfade unterscheiden sich nur hinsichtlich der Fernsehdauer. Die Form der Freizeitgestaltung und die Fokussierung auf *gemischte Unterhaltungssendungen* stimmen dagegen überein. Bei gleichem Fernsehprogramm und einer *aktiv/kreativ* genutzten sonstigen Freizeit scheint auch eine *mittlere* Fernsehdauer im Sinne überdurchschnittlicher Mathematikleistungen vertretbar zu sein. Dennoch gilt es zu beachten, dass die *Mittelsehergruppe* lediglich 5,6 % der Stichprobe repräsentiert, während die *Wenigsehergruppe* mehr als doppelt so viele Probanden beinhaltet und damit in seiner Aussagekraft mehr Gewicht und Bedeutung besitzt.

Wie schnell *Mittelseher* mit einem *gemischten Unterhaltungsprogramm* im Leistungsniveau abfallen, wenn an die Stelle der *aktiv/kreativen* Freizeitbeschäftigung eine *passiv/kognitiv anspruchslose* Form tritt, wird im Folgenden offensichtlich. Diese Probanden erhalten bereits leicht unterdurchschnittliche Beurteilungen in beiden mathematischen Dimensionen, so dass die Wichtigkeit des Freizeitverhaltens deutlich wird. Dies gilt sogar dann, wenn die Probanden *Wettkampfs- und Bildungssendungen* statt der gemischten Unterhaltung verfolgen. Der Konsum solcher Sendungen führt unter den Umständen eines *mittellangen* Fernsehkonsums und einem *anspruchslosen* Zeitvertreib in der Freizeit zu keinerlei (positiven) Zusammenhängen mit der Mathematikleistung. *Mittelseher* erhalten folgerichtig nur in Verbindung mit

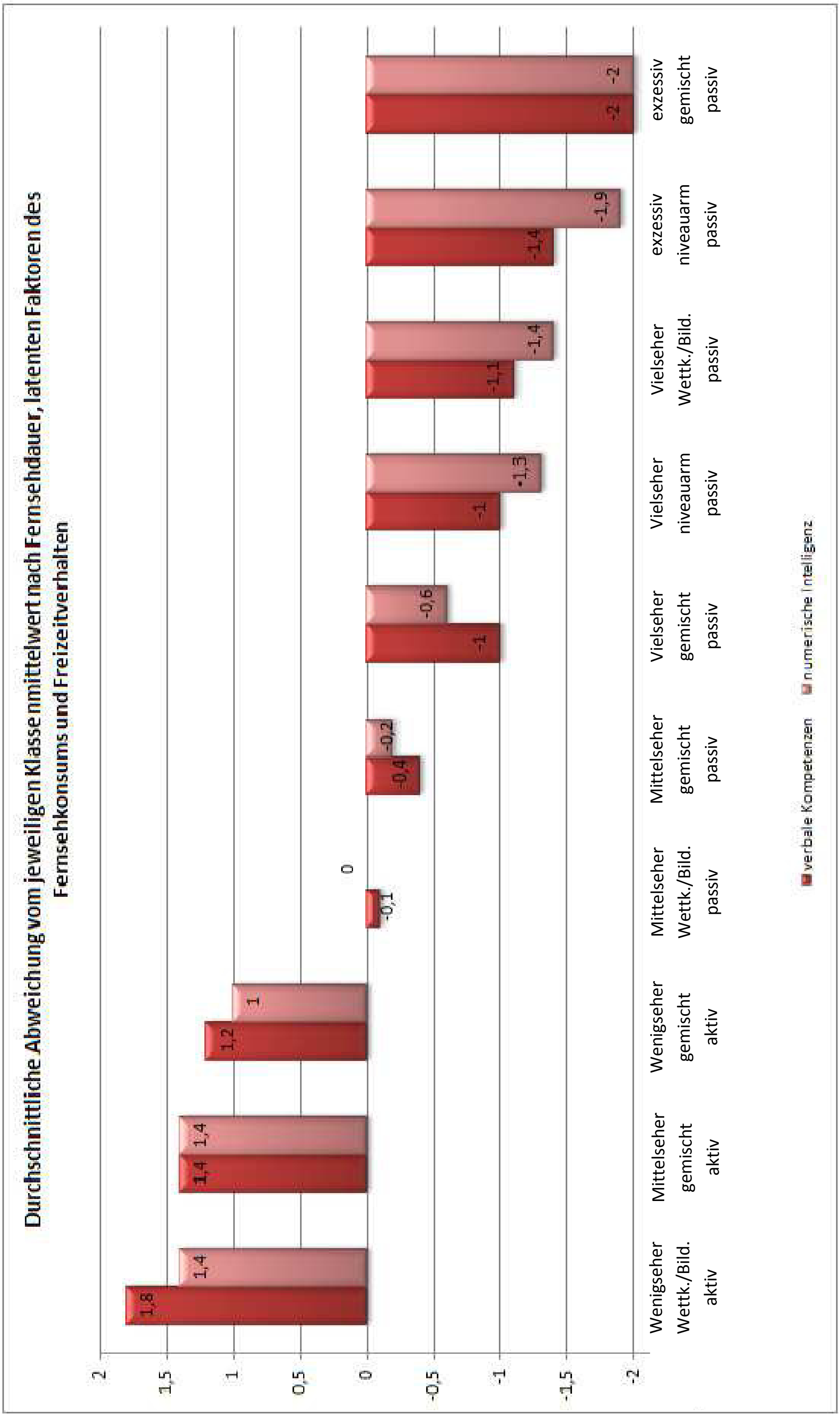


Abbildung 9.3 Durchschnittliche Abweichung vom jeweiligen Klassenmittelwert nach Fernsehdauer, latenten Faktoren des Fernsehkonsums und Freizeitverhalten

einem *aktiv/kreativen* Freizeitverhalten überdurchschnittliche Lehrerurteile. Die konsumierten Sendungen spielen dabei eine untergeordnete Rolle. Zwar sind bei *passiv/kognitiv anspruchslosen Mittelsehern* auch negative Zusammenhänge nur unter dem Mikroskop zu erkennen, doch ist die Differenz nach oben enorm.

Die fünf *Vielseher*- und *Exzessivseherpfade* stehen unabhängig von den latenten Faktoren des Fernsehkonsums generell in einem negativen Zusammenhang zu den Mathematikleistungen. Diese Beziehung wird zusätzlich dadurch verstärkt, dass diese Probanden grundsätzlich lieber *passiv/kognitiv anspruchslosen* Freizeitaktivitäten nachgehen. Bei einem solchen Verhalten verliert das Fernsehprogramm vollkommen an Bedeutung, so dass die *Vielsehergruppe* mit *niveauarmen* Programminhalten sich sogar knapp vor der *Vielsehergruppe* einsortiert, die Sendungen des Faktors *Wettkampf und Bildung* bevorzugt. Die Differenz beträgt zwar sowohl bei den *verbalen Kompetenzen* als auch bei der *numerischen Intelligenz* lediglich 0,1 Bewertungspunkte, ist aber dennoch vorhanden. Insgesamt liegen die *Vielseherpfade* zwischen **0,6** und **1,4** Bewertungspunkten unterhalb des jeweiligen Klassenmittelwerts. Bei den *passiven Exzessivsehern* besitzt das konsumierte Fernsehprogramm ebenfalls nur eine geringe Bedeutung. In diesem Fall bedeutet dies, dass die *passiven Exzessivseher* mit *gemischter Unterhaltung* (**-2,0**) noch ein klein wenig schlechter bewertet werden, als die ein *niveauarmes* Programm präferierenden *passiven Exzessivseher* (verbal: **-1,4**; numerisch: **-1,9**).

Die in der Vermutung gemachte Annahme wird durch die Daten im Diagramm in vollem Umfang bekräftigt, wobei der Fernsehdauer und dem Freizeitverhalten die Hauptrollen zukommen, während die latenten Faktoren des Fernsehkonsums lediglich die Nebenrolle übernehmen. Gerade bei einer erhöhten Fernsehdauer und *passiv/kognitiv anspruchslosen* Hobbys ist das Fernsehprogramm zu vernachlässigen. Wird *wenig* ferngesehen und ein *aktiv/kreatives* Freizeitprogramm realisiert, können *Wettkampfs- und Bildungssendungen* den positiven Effekt zusätzlich intensivieren.

Der bereits isoliert festgestellte Zusammenhang zwischen den drei Variablen und den Mathematikleistungen wird durch die durchgeführte Kombination noch weiter verstärkt!

Eine weitere Erkenntnis, die Abbildung 9.1 und 9.3 ermöglichen, ist die Folgende: Bei den überdurchschnittlich bewerteten Pfaden werden vor allem die *verbalen Kompetenzen* besonders gut bewertet, während die deutlich unterdurchschnittlich beurteilten Gruppen besonders im *numerischen* Bereich schwächeln. Eine mögliche Interpretation ist, dass das Fernsehen als visuelles und akustisches Medium gerade den schwächeren Schülerinnen und Schülern immerhin ein wenig *deutsche* Sprache vermittelt, wodurch der negative Einfluss auf die *numeri-*

sche Intelligenz etwas stärker ausgeprägt ist. Die leistungsstärkeren Jugendlichen benötigen eine solche Unterstützung nicht.

Um zu verdeutlichen, dass die durchschnittlichen Abweichungen vom jeweiligen Klassenmittelwert unabhängig von etwaigen Ausreißern die oben angestellte Ergebnisinterpretation rechtfertigen, wird im Folgenden der prozentuale Anteil überdurchschnittlich bewerteter Probanden für drei ausgewählte Pfade dargestellt. Dabei handelt es sich um je eine *Wettkampfs- und Bildungssendungen* favorisierende *Wenig*-, *Mittel*- und *Vielsehergruppe*. Einzig die *Wenigseher* verhalten sich in der Freizeit *aktiv/kreativ*. Durch den Konsum der augenscheinlich niveauvollsten Fernsehsendungen ist der Vergleich dieser drei Pfade besonders interessant:

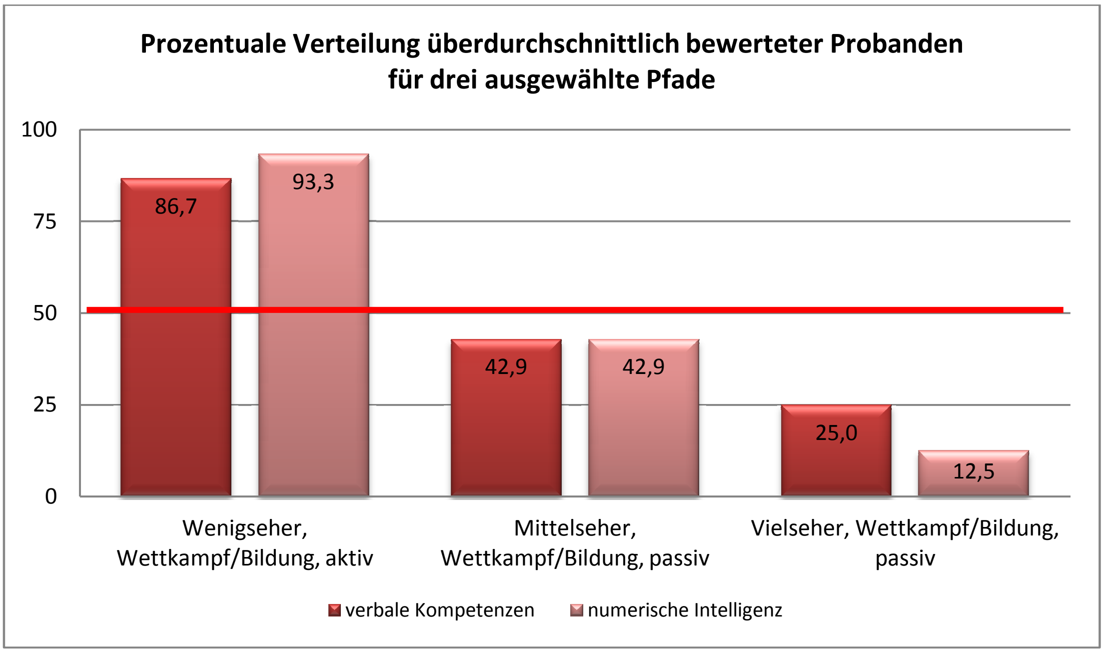


Abbildung 9.4 Prozentuale Verteilung überdurchschnittlich bewerteter Probanden für drei ausgewählte Pfade

Die Probanden des ersten Pfades weichen sowohl bei der Beurteilung der *verbalen Kompetenzen* als auch bei der Bewertung der *numerischen Intelligenz* am stärksten nach oben ab. Dieser Pfad steht also in einem positiven Zusammenhang zu den Mathematikleistungen. Abbildung 9.4 bestätigt dieses Ergebnis auch aus der Perspektive der prozentualen Anzahl überdurchschnittlich beurteilter Schülerinnen und Schüler. 86,7 % (*verbal*) bzw. 93,3 % (*numerisch*) der Probanden erhalten überdurchschnittliche Lehrerurteile. Die Kombination aus einer geringen Fernsehdauer, einem ansprechenden Fernsehprogramm und einer *aktiv/kreativen* Freizeitbeschäftigung scheint im Sinne der Mathematikleistungen die beste mediale Verhaltensweise zu sein. Verlängert sich jedoch die Fernsehdauer und verändert sich gleichzeitig das Freizeitverhalten zum negativen, können auch *Wettkampfs- und Bildungssendungen* einen negativen Zusammenhang nicht verhindern. Zum einen wird die entsprechende *Mittelsehergruppe* im Durchschnitt leicht negativ bewertet. Zum anderen erhalten lediglich 42,9 % eine überdurchschnittliche Bewertung ihrer mathematischen Fähigkeiten. Steigt die Fernsehdauer noch weiter an, werden je nach mathematischer Dimension bis zu 87,5 % (!) der Probanden unterdurchschnittlich beurteilt. Die obigen Zusammenhänge gelten dementsprechend unabhängig vom wenig robusten arithmetischen Mittelwert.

9.4 Die Hauptrollen „Fernsehdauer" und „Freizeitverhalten"

Sowohl die Ergebnisse aus der isolierten Auswertung der drei erhobenen Variablen als auch die bisherigen Tendenzen der kombinierten Methoden weisen vor allem die Fernsehdauer und das Freizeitverhalten als dominierende und einflussreichste Variablen aus.

Besonders deutlich wird dies anhand der Abbildungen 9.5 und 9.6:

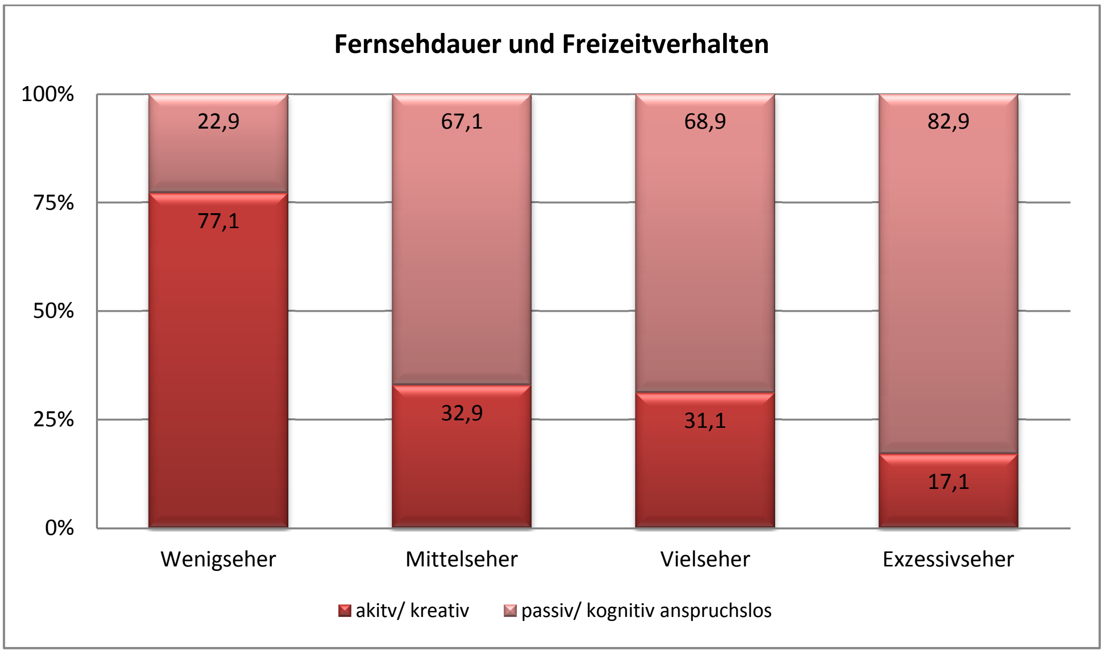


Abbildung 9.5 Fernsehdauer und Freizeitverhalten

Das Diagramm in Abbildung 9.5 zeigt zunächst die prozentuale Verteilung des bevorzugten Freizeitverhaltens in Abhängigkeit von der durchschnittlichen täglichen Fernsehdauer. Während ein Großteil der *Wenigseher aktiv* und *kreativ* ist (77,1 %), gilt dies bereits bei den *Mittelsehern* nur noch für 32,9 % der Probanden. Die *Vielseher* (68,9 %) und insbesondere die *Exzessivseher* (82,9 %) beschäftigen sich nahezu ausschließlich *passiv* und *kognitiv anspruchslos*. Verbringen Jugendliche einen Großteil des Tages vor dem Fernseher, bleibt keine Zeit für anspruchsvolle und möglicherweise zeitintensive Hobbys. Man denke dabei an das Lernen eines Musikinstruments, das Lesen eines dicken Buches oder stundenlanges Training im Sportverein. Die *passiven* Beschäftigungen wie Internet, chatten oder das Spielen mit dem Smartphone sind problemlos auch während des Fernsehkonsums möglich.

Dieses Verhalten äußert sich auch in den Mathematikleistungen, wie das Diagramm in Abbildung 9.6 unterstreicht. *Aktive Wenig-* und *Mittelseher* weichen in beiden mathematischen Dimensionen deutlich nach oben ab. Bei einer sinnvollen sonstigen Freizeitbeschäftigung ist auch ein *mittellanger* Fernsehkonsum hinsichtlich der Mathematikleistungen als legitim zu bezeichnen. Zwischen *passiven Viel-* sowie *Exzessivsehern* und den mathematischen Leistungen besteht dahingegen ein negativer Zusammenhang.

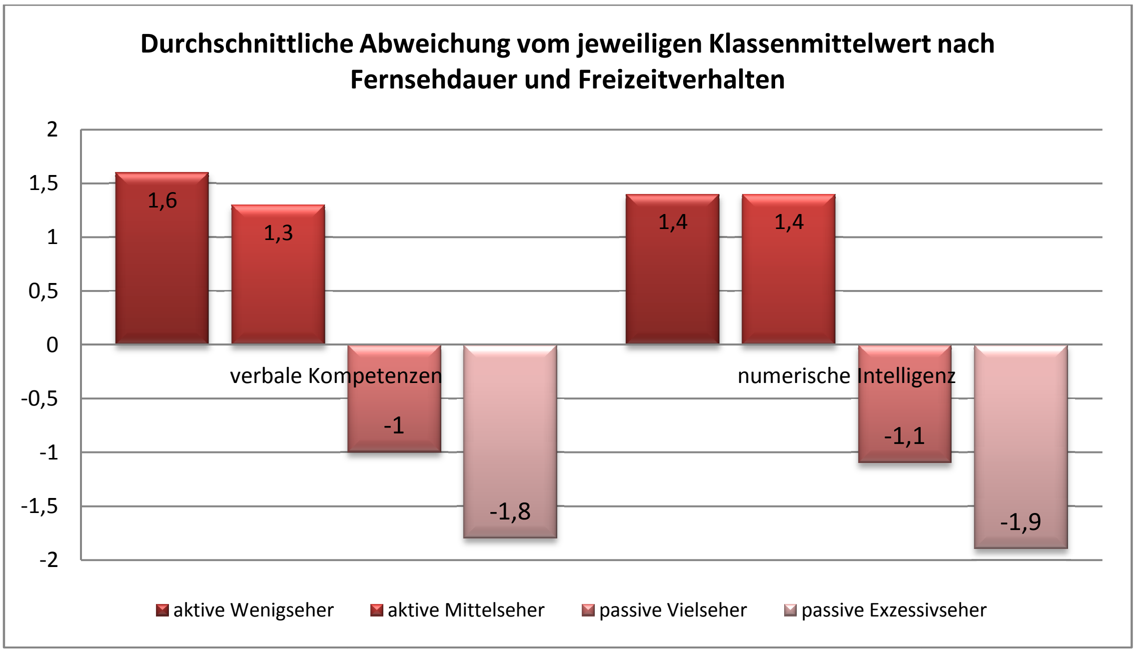


Abbildung 9.6 Durchschnittliche Abweichung vom jeweiligen Klassenmittelwert nach Fernsehdauer und Freizeitverhalten

9.5 Die Nebenrolle „latente Faktoren des Fernsehkonsums"

Derartige extrem ausgebildete Beziehungen bestehen zwischen den latenten Faktoren des Fernsehkonsums und der Fernsehdauer bzw. dem Freizeitverhalten nicht. Zwar kann die Art des Fernsehkonsums letzte Richtungsänderungen bzw. Effektverstärkungen begründen, da die Probanden ihr Fernsehprogramm häufig entsprechend der anderen beiden Variablen auswählen, doch zeichnen sich die Programminhalte zur Enttäuschung zahlreicher Fernsehbefürworter kaum als zentrales Element bei der Erklärung der identifizierten Zusammenhänge aus. Dies wurde bereits im Rahmen der isolierten Auswertung deutlich. Die gemischte *Unterhaltung* sowie der Faktor *Wettkampf und Bildung* stehen insgesamt weder in einem positiven noch in einem negativen Zusammenhang zu den Mathematikleistungen der Probanden. Lediglich die *niveauarme Unterhaltung* steht in einer negativen Beziehung zu den Lehrerurteilen. Kombiniert man nun die latenten Faktoren des Fernsehkonsums einzeln mit den anderen beiden Variablen, wird die Nebenrolle ebenfalls ersichtlich:

In Abbildung 9.7 wird eine solche zweier-Kombination zunächst für die Fernsehdauer und die latenten Faktoren des Fernsehkonsums gebildet und zeigt jeweils die durchschnittliche Abweichung vom Klassenmittelwert für die *verbalen Kompetenzen*. Das Ergebnis ist analog auf die *numerische Intelligenz* übertragbar.

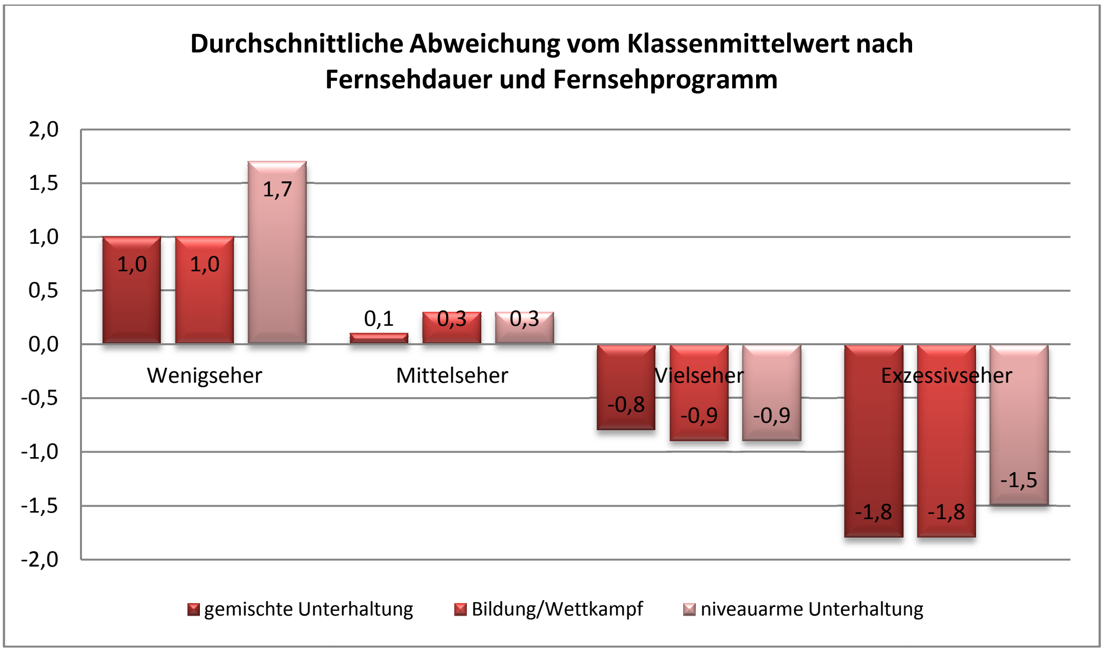


Abbildung 9.7 Durchschnittliche Abweichung vom Klassenmittelwert nach Fernsehdauer und Fernsehprogramm

Die *Wenigseher* werden allesamt überdurchschnittlich bewertet, insbesondere gilt dies sogar für diejenigen Jugendlichen, die *niveauarme* Fernsehsendungen bevorzugen. Bei den *Mittelsehern* ist der positive Zusammenhang nur noch minimal, aber weiterhin nicht abhängig von den konsumierten Programminhalten. Sämtliche *Vielsehergruppen* werden ebenfalls ähnlich beurteilt und liegen dicht beieinander. Die Gruppen der *exzessiv* fernsehenden Schülerinnen und Schüler macht ebenfalls keine Ausnahme. Geschlossen werden diese Jugendlichen unterdurchschnittlich und nahezu identisch beurteilt; unabhängig vom gewählten latenten Faktor.

Eine ähnliche Feststellung ermöglicht auch Abbildung 9.8, in welcher die *numerische Intelligenz* der Probanden in Abhängigkeit vom Fernsehprogramm und dem Freizeitverhalten illustriert wird. Offenkundig hängt auch in dieser zweier-Kombination eine positive bzw. negative Abweichung vom jeweiligen Klassenmittelwert weniger vom Fernsehprogramm als vom Freizeitverhalten ab. Die *aktiv/kreativen* Probanden erzielen unabhängig von den latenten Faktoren des Fernsehkonsums überdurchschnittliche Bewertungen, während die *passiv/kognitiv anspruchslosen* Jugendlichen generell mit einer negativen Abweichung rechnen müssen. Lediglich die Stärke des Effekts wird von den Programminhalten des Fernsehkonsums mitbestimmt, wobei *Wettkampfs- und Bildungssendungen* weder positiv noch negativ herausstechen.

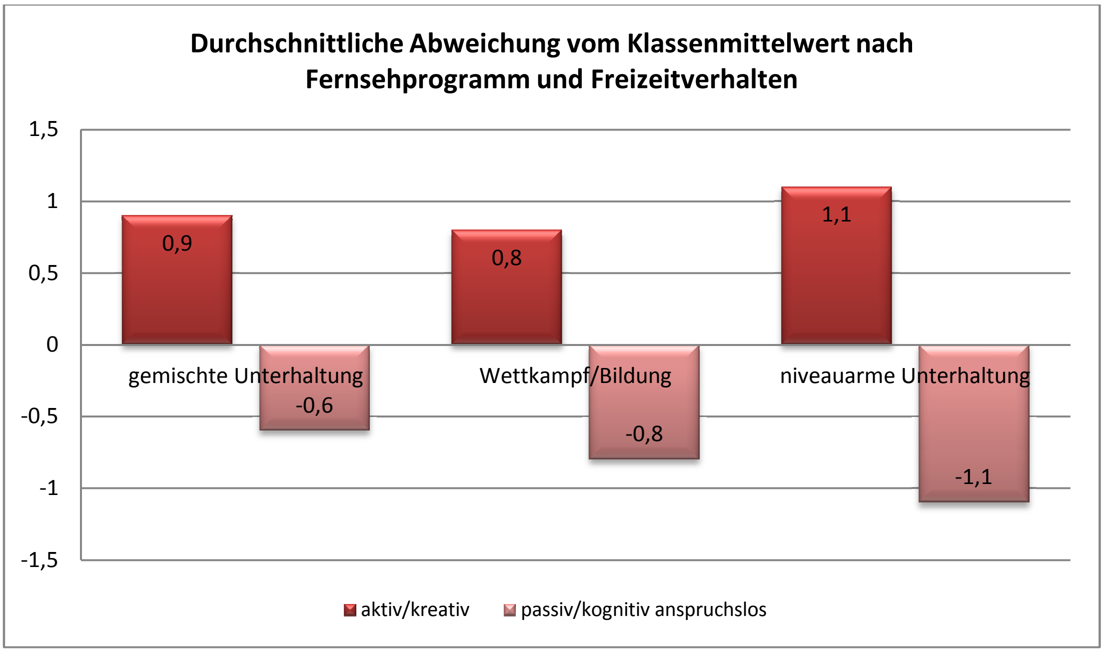


Abbildung 9.8 Durchschnittliche Abweichung vom Klassenmittelwert nach Fernsehprogramm und Freizeitverhalten

In den Abschnitten 9.4 und 9.5 konnte anhand von zweier-Kombinationen die Vermutung belegt werden, dass die identifizierten Beziehungen zwischen den erhobenen Variablen und den Mathematikleistungen der Probanden vor allem von der Fernsehdauer und dem Freizeitverhalten abhängen. Tatsächlich spielen die latenten Faktoren des Fernsehkonsums nur eine Nebenrolle!

9.6 Kombinierte geschlechtsspezifische Auswertung

Musste ein geschlechtsspezifischer Unterschied in der Beziehung zwischen dem Fernseh- bzw. dem Freizeitverhalten und den Mathematikleistungen von Jungen und Mädchen im Verlauf der isolierten Auswertung meist spitzfindig gesucht werden, ergibt sich bei der Kombination der drei Variablen vor allem in der Zusammensetzung der Pfade eine erhebliche Differenz zwischen den Geschlechtern. Dazu betrachte man die unten stehende Tabelle 9.2, in der die jeweils höchstfrequentiertesten Gruppen abgebildet werden.

Während sich die Mädchen gleichmäßig auf die Pfade verteilen, ist bereits ein Drittel der männlichen Probanden in den beiden ersten Gruppen enthalten. Erst die weiteren männlichen Pfade werden durch in etwa gleich große Gruppen gebildet. Dies deutet darauf hin, dass die Homogenität im Fernseh- und Freizeitverhalten bei den Jungen wesentlich größer ausgeprägt ist, als bei den Mädchen. Diese These findet zusätzliche Unterstützung, da alle bedeutenden männlichen Kombinationen den gleichen latenten Faktor des Fernsehkonsums beheimaten: Die *gemischte Unterhaltung* dominiert unabhängig von der Fernsehdauer und dem Freizeitverhalten. Bei den Mädchen treten im Gegensatz dazu drei verschiedene Faktoren in Erscheinung. Zwar übernehmen *Wettkampfs- und Bildungssendungen* die Führungsrolle, doch spielt die *niveauarme Unterhaltung* gerade bei den *Viel*- und *Exzessivseherinnen* eine wichtige Rolle.

Tabelle 9.2

Die wichtigsten kombinierten Gruppen nach Geschlecht

weiblich			männlich
Wenigseher, gemischte Unterhaltung, aktiv	9,4 %	17,5 %	Mittelseher, gemischte Unterhaltung, passiv
Wenigseher, Wettkampf/Bildung, aktiv	8,5 %	15,1 %	Wenigseher, gemischte Unterhaltung, aktiv
Mittelseher, Wettkampf/Bildung, passiv	8,5 %	7,9 %	Mittelseher, gemischte Unterhaltung, aktiv
Vielseher, Wettkampf/Bildung, passiv	8,5 %	7,1 %	Exzessivseher, gemischte Unterhaltung, passiv
Exzessivseher, Wettkampf/Bildung, passiv	8,5 %	7,1 %	Vielseher, gemischte Unterhaltung, passiv
Mittelseher, gemischte Unterhaltung, passiv	7,5 %		
Vielseher, niveauarme Unterhaltung, passiv	7,5 %		

Die Zusammenhänge zwischen den *verbalen* und *numerischen* Lehrerurteilen und den wichtigsten Gruppen sind allerdings bei beiden Geschlechtern ein ziemlich getreues Abbild der gesamten Stichprobe. Die Abweichungen sind eher minimaler Natur. Bei den Jungen erhalten die *aktiv/kreativen Wenig*- (*verbal* : **1,1**; *numerisch*: **1,0**) und *Mittelseher* (*verbal*: **1,3**; *numerisch*: **1,3**) die im Vergleich zum (männlichen) Klassenmittelwert besten Bewertungen. Die *passiven Mittel*- (*verbal*: **-0,7**; *numerisch*: **-0,4**), *Viel*- (*verbal*: **-1,2**; *numerisch*: **-0,8**) und *Exzessivseher* (*verbal*: **-2,4**; *numerisch*: **-2,2**) werden sukzessive schlechter beurteilt. Die Bedeutung der beiden dominanten Variablen „Fernsehdauer" und „Freizeitverhalten" wird dementsprechend auch an dieser Stelle deutlich. Positiv weichen nur diejenigen Jungen ab, die eine gemäßigte Fernsehdauer <u>und</u> ein anspruchsvolles Freizeitprogramm vorweisen können.

Ein ähnliches Bild ergibt sich auch bei den weiblichen Probanden. Die *aktiv/kreativen Wenigseherinnen* erhalten in beiden mathematischen Dimensionen die besten Lehrerurteile. Den Ausschlag über die am besten bewertete Kombination gibt dabei das vorwiegend konsumierte Fernsehprogramm, so dass sich *Wettkampf und Bildung* (*verbal*: **2,1**; *numerisch*, **2,0**) vor der *gemischten Unterhaltung* (*verbal*: 0,9; *numerisch*: 0,5) einreiht. Es folgen die beiden Cluster der *passiven Mittelseherinnen*, wobei erneut der Faktor *Wettkampf und Bildung* mit den etwas besseren Bewertungen in Verbindung steht (*verbal*: **-0,1** zu **-0,2**; *numerisch*: **0,4** zu **-0,4**). Auf den letzten drei Plätzen rangieren konsequenterweise die *passiven Viel*- und *Exzessivseherinnen.* Interessanterweise befinden sich die mathematischen Kompetenzen der *passiven Exzessivseherinnen* trotz eines vorwiegend *niveauarmen Fernsehprogramms* in einem weniger stark negativen Zusammenhang, als dies bei den Leistungen der männlichen *Exzessivseher* der Fall ist. Der Faktor *Wettkampf und Bildung* scheint bei den weiblichen

Probanden nach der Kombination der einzelnen Variablen an Bedeutung gewonnen zu haben, ist er im Vergleich zur *gemischten Unterhaltung* mehrmals mit einer positiven Abweichung verbunden.

Eine Prognose darüber zu treffen, welches geschlechtsspezifische Fernseh- und Freizeitverhalten empfehlenswerter ist, ist nur schwer möglich. Auf Basis des Vergleichs der *aktiven Wenig-* und *Mittelsehergruppen* ist allerdings zumindest eine tendenzielle Aussage möglich:

Tabelle 9.3

Vergleich zwischen Mädchen und Jungen

	weiblich		männlich	
	Anteil	verbal/ numerisch	Anteil	verbal/ numerisch
aktiv/kreative Wenig- und Mittelseher	29,2 %	1,6 / 1,4	38,1 %	1,4 / 1,2
Rest	70,8 %	-0,7 / -0,6	61,9 %	-0,8 / -0,8

Wie die Daten in Tabelle 9.3 unterstreichen, steht ein *aktives* Freizeitverhalten in Kombination mit einer maximal *mittellangen* Fernsehdauer in einem positiven Zusammenhang zu den mathematischen Leistungen. Demnach verhalten sich, im Gegensatz zu anderslautenden Meinungen, die männlichen Probanden in ihrer schulfreien Zeit erfolgsversprechender: Zwar ist der Effekt für beide Geschlechter recht ähnlich, wobei die weiblichen Jugendlichen sogar unwesentlich besser abschneiden, doch bevorzugen viel mehr Jungen (über 38 %) als Mädchen (unter 30 %) ein solch positives Verhalten.

9.7 Kombinierte schulformspezifische Auswertung

Zum Abschluss der Auswertung bleibt letztlich die Betrachtung der möglichen Zusammenhänge zwischen den kombinierten Variablen und den Mathematikleistungen in Abhängigkeit von der Schulformzugehörigkeit der Probanden übrig. In Anlehnung an die geschlechtsspezifische Ergebnisanalyse konzentriere man sich in einem ersten Schritt auf die für jede Schulform typischen Pfade des Fernseh- und Freizeitverhaltens, dargestellt in Tabelle 9.4:

Tabelle 9.4

Die wichtigsten kombinierten Gruppen nach Schulform

Gymnasium	Gesamtschule	Hauptschule
wenig, gemischt, aktiv (28,9 %) mittel, gemischt, passiv (12,2 %) mittel, gemischt, aktiv (8,9 %) wenig, Wettk./Bild., aktiv (7,8 %)	mittel, gemischt, passiv (18,3 %) viel, Wettk./Bild., passiv (9,6 %) mittel, Wettk./Bild., passiv (8,7 %) exzessiv, niveauarm, passiv (7,7 %) wenig, Wettk./Bild, aktiv (6,7 %) viel, gemischt, passiv (6,7 %)	exzessiv, Humor, passiv (10,7 %) exzessiv, niveauarm, passiv (10,7 %) exzessiv, gemischt, passiv (10,7 %) viel, gemischt, aktiv (10,7 %) wenig, gemischt, aktiv (10,7 %) wenig, Musik, aktiv (10,7 %)

Offensichtlich verhalten sich die ein Gymnasium besuchenden Schülerinnen und Schüler in ihrem Fernseh- und Freizeitverhalten ziemlich homogen und vorbildlich. Fast 30 % der Gymnasiasten schauen *wenig* fern, konsumieren *gemischte Unterhaltungssendungen* und beschäftigen sich in ihrer Freizeit *aktiv/kreativ*. Auch die nachfolgenden Gruppen werden von *Wenig*- und *Mittelsehern* gebildet, die kein *niveauarmes* Fernsehen bevorzugen und sich in den meisten Fällen *aktiv/kreativen* Freizeitaktivitäten widmen. Die Mathematikleistungen dieser Probanden liegen durchschnittlich in beiden mathematischen Dimensionen einen Bewertungspunkt oberhalb des Klassendurchschnitts. Bereits diese Feststellung bestätigt gängige Vorurteile, lässt Leistungsunterschiede vermuten und trennt die Gymnasiasten damit von den Jugendlichen der anderen beiden Schulformen.

An der Gesamtschule lässt diese Homogenität nach. Die hochfrequentierteste Gruppe (18,3 %) ist zudem eine Kombination, die *Mittelseher* und *passiv/kognitiv anspruchslose* Probanden enthält. Gerade das *passive* Freizeitverhalten steht wie erklärt in einem negativen Zusammenhang zu den Mathematikleistungen, so dass es nicht verwundert, dass diese Jugendlichen nach unten abweichende Lehrerbewertungen erhalten. Das größte Cluster, welches *Wenigseher* und *aktiv/kreative* Schülerinnen und Schüler beinhaltet, liegt deutlich oberhalb des Klassenmittelwerts (*verbal*: **2,0**; *numerisch*: **1,**7). Leider repräsentieren diese Jugendlichen lediglich 6,7 % der Gesamtschulstichprobe. Dementsprechend kann das Fernseh- und Freizeitverhalten als weniger intellektuell anspruchsvoll beschrieben werden, als dies am Gymnasium der Fall ist.

Die am Gymnasium identifizierte Homogenität weicht an der Hauptschule einer vollkommen heterogenen Schülerschaft. Keine Gruppe schließt mehr als 10,7 % der Jugendlichen ein. Dabei sind sowohl vorbildliche *Wenigsehergruppen* mit *aktiver* Freizeitgestaltung als auch kritisch zu bewertende *Exzessivsehercluster* mit *niveauarmen* Programminhalten vorhanden. In Abhängigkeit von der jeweiligen Gruppenzugehörigkeit weichen die Lehrerurteile zwischen **3,0** Bewertungspunkten nach oben und **1,9** Bewertungspunkten nach unten ab.

Zusammenfassend kann festgehalten werden, dass sich viele Gymnasiasten normgerecht verhalten, während die Homogenität als auch der Anspruch des Freizeit- und Fernsehverhaltens von der Gesamtschule bis hin zur Hauptschule immer weiter nachlässt. Dementsprechend überrascht es auch nicht, dass Gymnasiasten in der Summe die mit Abstand besten Lehrerurteile ihrer mathematischen Leistungen erhalten. Dazu betrachte man abschließend Tabelle 9.5:

Tabelle 9.5

Vergleich zwischen den Schulformen

	Gymnasium		Gesamtschule		Hauptschule	
	Anteil	verbal/ numerisch	Anteil	verbal/ numerisch	Anteil	verbal/ numerisch
aktiv/kreative Wenig- und Mittelseher	54,4 %	1,2 / 1,0	20,2 %	1,8 / 1,7	23,7 %	2,6 / 2,6
Rest	45,6 %	-1,4 / -1,2	79,8 %	-0,5 / -0,4	76,3 %	-0,8 / -0,8

Über die Hälfte der Gymnasiasten schauen maximal *mittellang* fern und beschäftigen sich in der sonstigen Freizeit *aktiv/kreativ*. Diese Probanden werden in beiden mathematischen Bereichen überdurchschnittlich bewertet. Die übrigen Jugendlichen erhalten ebenso deutliche unterdurchschnittliche Bewertungen. An der Gesamt- und Hauptschule gehören maximal 24 % zu den vorbildlichen *aktiv/kreativen Wenig-* und *Mittelsehern*. Zwar weichen diese Jugendlichen im Durchschnitt noch stärker nach oben ab, als es die Gymnasiasten tun, doch trifft dieser Effekt eben nur auf die absolute Minderheit zu. Die absolute Mehrheit der Schülerinnen und Schüler von den Gesamt- und Hauptschulen befinden sich dagegen auf einem mathematischen Wissensstand, der von den entsprechenden Lehrerinnen und Lehrern im Mittel als unterdurchschnittlich beschrieben wird.

9.8 Zusammenfassung

Die bei der isolierten Auswertung identifizierten Zusammenhänge bleiben auch nach der Kombination erhalten. Der Schulerfolg ist in Mathematik dabei vor allem mit der Freizeitbeschäftigung und der Fernsehdauer verbunden. Wer sich *aktiv/kreativ* verhält und nur maximal *mittellang* fern sieht, hat optimale mediale Voraussetzungen, gute Mathematikleistungen zu entwickeln. Daran kann selbst ein durchschnittlich anspruchsvolles Fernsehprogramm wenig ändern. Unglücklicherweise verhält sich die Mehrheit der Probanden nicht derart wünschenswert. Viele Jugendliche beschäftigen sich nämlich lieber *passiv* bzw. *kognitiv an-*

spruchslos. In Kombination mit einer *langen* oder *exzessiven* Fernsehdauer sind offensichtliche negative Zusammenhänge zu den mathematischen Leistungen feststellbar. Auch der Konsum von *Wettkampfs- und Bildungssendungen* reicht in einem solchen Fall maximal zur Schadensbegrenzung. Dementsprechend können den Variablen „Fernsehdauer" und „Freizeitverhalten" die Hauptrollen zugeschrieben werden, während den „latenten Faktoren des Fernsehkonsums" endgültig lediglich die Nebenrolle bleibt.

Der beobachtete Zusammenhang bleibt geschlechts- und schulformunabhängig bestehen. Allerdings konnte festgestellt werden, dass mehr Jungen als Mädchen bzw. mehr Gymnasiasten als Schülerinnen und Schüler der Gesamt- und Hauptschule ein normgerechtes und vorbildliches Fernseh- und Freizeitverhalten zeigen. Dementsprechend werden die Mathematikleistungen der Jungen und der Gymnasiasten häufiger positiv bewertet. Mädchen und Jugendliche der Gesamt- und Hauptschule gehören eher zur Risikogruppe.

10. Fazit und Diskussion

Im Verlauf dieser Dissertation wurde das Ziel verfolgt, eine empirisch und methodisch fundierte Antwort auf die viel diskutierte Frage zu geben, ob die Nutzung des visuellen und akustischen Mediums Fernsehen in einem Zusammenhang zu Schulleistungen von Schülerinnen und Schülern steht. Da die Vermutung aufgestellt wurde, dass eine globale Antwort aufgrund der Wirkungsmechanismen des Fernsehens a priori nicht auszusprechen ist, wurde in dieser Arbeit eine lokale Untersuchung, die numerischen bzw. mathematischen Leistungen der Probanden betreffend, vorgeschoben.

Im Allgemeinen Teil konnte anhand repräsentativer Studien des Medienpädagogischen Forschungsverbunds Südwest dazu gezeigt werden, dass das Fernsehen im Alltag vieler Kinder, Jugendlicher und Familien fest verankert ist. Trotz einer neuen Konkurrenzsituation durch Smartphones, Computer und die unendlichen Weiten des Internets behauptet sich das Fernsehen nach wie vor als wichtiger und viel genutzter medialer Zeitvertreib. Vor allem im Kindesalter ist das Medium sehr dominant, bei den Jugendlichen wird die Nutzung immer autonomer und in den Familien gehören Themen, die in Zusammenhang zu dem Fernsehen stehen, zur Basis vieler Gespräche. Das Fernsehen fasziniert Heranwachsende, weil es ihnen ermöglicht in Sekundenschnelle dem Alltag zu entfliehen und im Gegensatz zu Büchern und Geschichten mit Farben, Geräuschen und Musik arbeitet. Diese Faktoren erleichtern es den Kindern und Jugendlichen in fremde Welten einzutauchen und ihre Tagträume zu erleben. Leider gehen dabei Fantasie und Vorstellungsvermögen verloren.

Im dritten Kapitel dieser Arbeit konnte durch die Aufführung verschiedener Meinungen und bestehender empirischer Studien bereits verdeutlicht werden, dass die Diskussion um Zusammenhänge zwischen Fernsehkonsum und Bildung von Schülerinnen und Schülern kontrovers geführt wird. Selbst Medienexperten und Wissenschaftler sind sich bei der Bewertung des Mediums uneins. Die kritischen Meinungen und Studien verweisen vor allem darauf, dass der überhöhte Medienkonsum, unangebrachte Fernsehsendungen sowie die emotionale Einstellung der Jugendlichen zu schwächeren Schulleistungen führen können. Studien des Kriminologischen Forschungsinstituts Niedersachsens und der *Mensch-Zeichen-Test* von Peter Winterstein (2006) scheinen negative Auswirkungen zu bestätigen. Auch Manfred Spitzer (2012) verweist auf weitere empirische Untersuchungen, die das Fernsehen teilweise in Misskredit bringen. Wissenschaftliche Studien, die sich mit pädagogisch geprägten Fernsehsendungen auseinandersetzen, erkennen aber auch das Potenzial, welches das Medium bietet: den Kindern und Jugendlichen Spaß und Freude am Lernen vermitteln! Es gibt nämlich durchaus Sendungen, von denen Kinder und Jugendliche etwas lernen können. Leider werden diese Sendungen allerdings nicht ausreichend geschaut und beworben. Stattdessen bevorzugen viele Jugendliche vor allem Serienformate von RTL und ProSieben. Letztlich sind sich aber alle Experten und Forscher darüber einig, dass der Medienerziehung im Elternhaus eine große Bedeutung zukommt. Die Eltern sind dafür verantwortlich, dass ihre Kinder die Medien verantwortungsvoll nutzen.

Das Herzstück dieser Dissertation ist der Methodische Teil, indem eine eigenständig durchgeführte Studie vorgestellt worden ist. Ziel dieser Studie war es vor allem, auf mathematisch und statistisch adäquate Weise zu untersuchen, ob zwischen Fernsehverhalten einerseits und in der Schule gemessenen Mathematikleistungen von Jugendlichen andererseits messbare Zusammenhänge bestehen und das Fernsehverhalten somit einen Grund für schwache Leistungen in Mathematik repräsentiert.

Zur Beantwortung der Forschungsfrage sind mehrere Fragebogenaktionen unter 232 Schülerinnen und Schülern der siebten Jahrgangsstufe an einem Gymnasium, zwei Gesamtschulen und einer Hauptschule durchgeführt worden, bei denen drei Teilbereiche des Fernseh- und Freizeitverhaltens abgefragt wurden. Dabei konnte gezeigt werden, dass die Fernsehdauer in vier Zeitintervalle (*Wenigseher*, *Mittelseher*, *Vielseher*, *Exzessivseher*) unterteilt werden kann, fünf latente Faktoren dem Fernsehverhalten der Jugendlichen zugrunde liegen (*gemischte Unterhaltung*, *Wettkampf und Bildung*, *niveauarme Unterhaltung*, *humorvolle Unterhaltung*, *musikalische Unterhaltung*) und das Freizeitverhalten der Probanden entweder als *aktiv/kreativ* oder aber als *passiv/kognitiv anspruchslos* beschrieben werden kann. Zudem wurde durch eine Lehrerumfrage die Mathematikleistung *messbar* gemacht. Demnach basieren gute mathematische Leistungen auf zwei zentralen Faktoren: der *verbalen Kompetenz* und der *numerischen Intelligenz*.

Eine wesentliche Stärke der Datenerhebung der fragebogenbasierten Untersuchung ist sicherlich, dass im Vergleich zu anderen Studien aus dem Themenfeld „Bildung und Fernsehen" die erhobenen Variablen gleichbedeutend sind. Dies meint, dass sowohl die Fernsehdauer als auch die Art des Fernsehprogramms berücksichtigt werden. Der *Mensch-Zeichen-Test* von Winterstein oder das *Sesamstraßenexperiment* haben ihre sich widersprechenden Forschungsergebnisse beispielsweise mit nur einer der beiden Variablen begründet. Damit wird in diesen Studien jeweils ein wichtiger Aspekt unterschlagen, welcher die gefunden Ergebnisse möglicherweise in einem anderen Licht präsentiert hätte.

Besonders hervorzuheben ist aber vor allem die komplexe Erhebung der latenten Faktoren des Fernsehverhaltens. Im Vergleich zu vielen anderen thematisch ähnlichen Untersuchungen, in denen entweder lediglich exemplarische Fernsehsendungen ausgewählt oder aber bestimmte Sehgewohnheiten als Antwortalternativen für die Probanden vorgegeben werden, basieren die Forschungsergebnisse in dieser Arbeit auf sämtlichen Fernsehsendungen, die im deutschen Fernsehen ausgestrahlt werden. Nicht der Verfasser, sondern die Probanden selbst haben aus der gigantischen Auswahl diejenigen Fernsehsendungen herausgefiltert, die in dieser Altersgruppe tatsächlich von Relevanz sind. So konnten die unterschiedlichen Sehgewohnheiten auf Basis von über 50 Fernsehsendungen ermittelt werden. Damit nimmt die durchgeführte Untersuchung sicherlich eine Vorreiterrolle ein. Ähnliches gilt zudem für die Extraktion der latenten Faktoren des Freizeitverhaltens.

Zudem ist herauszustellen, dass die verschiedenen Gruppen der Fernsehdauer (*Wenig-*, *Mittel-*, *Viel-* und *Exzessivseher*) nicht vom Verfasser vorgegeben wurden, sondern ebenfalls methodisch und statistisch fundiert herausgearbeitet worden sind. Lediglich mit Hilfe datenanalytischer Verfahren ist es möglich, objektiv festzustellen, welche Probanden im Vergleich zur gesamten Stichprobe wenig oder viel Zeit vor dem Fernseher verbringen und wie viele verschiedene Gruppen tatsächlich nötig sind, um die durchschnittliche Dauer des täglichen Fernsehkonsums signifikant zu beschreiben. Wichtig ist dabei, dass den Probanden keine Zeitspannen zum Ankreuzen vorgegeben wurden. Ein solches Untersuchungsdesign würde die nachfolgende Auswertung extrem einschränken.

Zugleich ist die Erhebung der Fernsehdauer aber auch ein problematischer Aspekt innerhalb der Studie, da die subjektiven Antworten der Probanden sicherlich von der tatsächlichen täglichen Fernsehdauer mehr oder weniger stark abweichen. In wenigen anderen empirischen Arbeiten (vgl. Würzburger Arbeitsgruppe) wird daher die Fernsehdauer beispielsweise über ein *Medientagebuch* erhoben, in welchem die Probanden ihren täglichen Fernsehkonsum exakt festhalten können. Allerdings bringt auch dieses Vorgehen Nachteile mit sich. Beispielsweise muss ein solches Tagebuch über einen längeren Zeitraum kontinuierlich und gewissenhaft geführt werden. Wird das Fernsehverhalten nur auf Basis der Tagebucheinträge eines Tages oder einer Woche erhoben, ist nicht auszuschließen, dass in dieser Woche aufgrund des möglicherweise schönen Wetters besonders wenig oder aber aufgrund des besonders spannenden Fernsehprogramms besonders viel ferngesehen wird. Erst bei einer langfristigen Untersuchung wird die Wirkung äußerer Einflüsse auf das Forschungsergebnis minimiert. Ob Probanden ein solch langfristiges Projekt gewissenhaft und motiviert umsetzen, ist zumindest fragwürdig. Daher erscheint es durchaus logisch, sich an dem Vorgehen des Medienpädagogischen Forschungsverbunds Südwest zu orientieren, die anhand von Selbsteinschätzungen seit Jahren erfolgreich die durchschnittliche Fernsehdauer von Jugendlichen erheben.

Rückblickend vermutet der Verfasser, dass es zudem interessant gewesen wäre, zu erheben, ob die Probanden ein eigenes Fernsehgerät in ihrem Zimmer besitzen. Diese Variable fehlt in der vorliegenden Arbeit leider.

Die größte Problematik im Verlauf der Datenerhebung ist aber vermutlich in der Messung der Mathematikleistung zu sehen. Sowohl die didaktische Diskussion als auch die mathematische Modellierung haben gezeigt, dass ein Konsens und allgemein akzeptierte Vorschläge zur Leistungsmessung noch nicht gefunden sind. Daher ist in dieser Arbeit eine Lehrerumfrage zur Leistungsmessung durchgeführt worden. Diese Art der Leistungsmessung ist weder vollkommen objektiv, noch entspricht sie dem, was Didaktiker gerne unter Leistungsmessung verstehen, doch wird diesem Mangel bereits im Titel der Arbeit Rechnung getragen, so dass die Einschätzung der Qualität der Leistungsmessung der Leserin bzw. dem Leser überlassen bleibt. Nicht zu unterschätzen ist allerdings, dass im schulischen Alltag nunmal die Lehrkräfte Schwierigkeiten und Fähigkeiten bei den einzelnen Schülerinnen und Schülern feststellen. Standardisierte Tests zur Bestätigung der Anfangsdiagnose werden, wenn überhaupt, erst im Anschluss an das Lehrerurteil durchgeführt. Daher wird in dieser Arbeit auf den Anfang der Leistungsmessung durch die Schule zurückgegriffen. Im Vergleich zur Notenvergabe eignet sich die Lehrerumfrage insofern, als dass die Lehrerinnen und Lehrer statt nur sechs Schulnoten eine stetige Bewertungsskala zur Leistungsbeurteilung zur Verfügung hatten. Darüber hinaus konnten sich die Lehrkräfte bei ihrem Urteil ausschließlich an den tatsächlichen Leistungen orientieren, statt Erziehungsziele o. ä. berücksichtigen zu müssen.

Um die Mathematikleistungen und das Fernsehverhalten in Bezug zu setzen, wurden in einem ersten Schritt die drei erhobenen Variablen (Fernsehdauer, latente Faktoren des Fernsehkonsums, latente Faktoren des Freizeitverhaltens) isoliert mit den Mathematikleistungen der Probanden verglichen. Es konnte festgestellt werden, dass vor allem die Fernsehdauer und das Freizeitverhalten in einem messbaren Zusammenhang zu den mathematischen Leistungen der Probanden stehen. Vor allem die *Wenigseher* und die *aktiv/kreativen* Jugendlichen erhalten überdurchschnittliche Lehrerurteile, während die *Viel-* und *Exzessivseher* sowie die *passiven* Probanden unterhalb der klasseninternen Bezugsnorm einsortiert werden. Der Zusammenhang zwischen Mathematikleistungen und den konsumierten Programminhalten ist wesentlich geringer, allerdings besteht zumindest eine negative Beziehung zwischen *niveauarmer* Fernsehunterhaltung und den zu vergleichenden Mathematikleistungen. Diese Zusammenhänge lassen sich insgesamt geschlechts- und schulformunabhängig beobachten, doch zeigt sich dass die Gesamt- und Hauptschulrezipienten wesentlich häufiger vom gewünschten Medienverhalten abweichen und somit öfter unterdurchschnittlich beurteilt werden. Schülerinnen und Schüler der Hauptschule, welche *wenig* fern sehen und sich zusätzlich *aktiv* oder *kreativ* beschäftigen, sind zwar eher die Ausnahme, erhalten aber im Vergleich zu den Klassenkameraden überaus positive Beurteilungen. Sich abweichend verhaltene Gymnasiasten erhalten besonders schlechte Bewertungen. Es scheint, als würde von den Gymnasiasten erwartet, dass sie sich normgerecht verhalten.

Abschließend wurden die drei Variablen miteinander kombiniert. Es konnte gezeigt werden, dass vor allem diejenigen Probanden, welche *aktiv/kreativen* Freizeitbeschäftigungen nachgehen und gleichzeitig *wenig* fernsehen die besten Lehrerurteile erreichen, während mit zunehmender Fernsehdauer und einem Wandel zu *passiv/kognitiv anspruchslosen* Freizeitaktivitäten die Schulleistungen in beiden mathematischen Dimensionen unter den Durchschnitt abrutschen. Die geschauten Fernsehsendungen spielen dabei lediglich eine Nebenrolle und können die Beobachtungen höchstens verstärken bzw. abschwächen. Auch nach der Kombination der Variablen unterstützt das Ergebnis viele von Manfred Spitzer zitierte Studien und Meinungen, nach denen alleine die Fernsehdauer, unabhängig von konsumierten Programminhalten, für schwache Schulleistungen mitverantwortlich ist. Wenn zu viel Zeit vor dem Fernseher verbracht wird, hilft auch die *Sesamstraße* nicht mehr. Zudem deutet die Tendenz darauf hin, dass mit zunehmender Fernsehdauer auch die Qualität der konsumierten Programminhalte nachlässt und eine hohe tägliche Fernsehdosis zusätzlich häufig *passive* und *kognitiv anspruchslose* Hobbys bedingt. Das Fernseh- und Freizeitverhalten von Jungen und Gymnasiasten ist am homogensten und entspricht zudem am ehesten den gesellschaftlichen Wunschvorstellungen. Das Verhalten von Gesamt- und Hauptschulprobanden ist wesentlich heterogener und zusätzlich seltener wünschenswert. Diese Probanden gehören zur Risikogruppe!

Insgesamt reiht sich die empirische Studie dieser Arbeit damit ein in die Menge der kritischen Studien, die dem Fernsehen vorwerfen in einem negativen Zusammenhang zu in der Schule erbrachten Leistungen von Kindern und Jugendlichen zu stehen. Wie bereits von Spitzer, dem KFN, Winterstein u. a. befürchtet, ist dabei vor allem eine überhöhte tägliche Fernsehdauer besonders problematisch. Jugendliche, die den Großteil der Freizeit vor dem Fernseher verbringen, verhalten sich zudem auch in der übrigen Zeit eher *passiv* und *kognitiv anspruchslos*. Daher verwundert es nicht, dass nicht nur die sprachlichen Leistungen in einem negativen Zusammenhang stehen, sondern nun auch gezeigt werden konnte, dass für den Fachbereich der Mathematik ähnliche Beobachtungen gelten. Nichtsdestotrotz ist es zwingend erforderlich festzuhalten, dass es sich dabei nur um ein Indiz handelt. Dies hat mehrere Ursachen:

(i) Einerseits bedingen sich die Fernsehdauer, die latenten Faktoren des Fernsehkonsums sowie das Freizeitverhalten gegenseitig, so dass hier Kausalketten nicht auszuschließen sind.

(ii) Darüber hinaus ist es nicht möglich, eine Einflussrichtung zu identifizieren. Ob schwache mathematische Leistungen zu einem überhöhten Fernsehkonsum führen oder ob das Fernsehverhalten für die gemessenen Mathematikleistungen verantwortlich ist, vermag die Studie nicht zu beantworten.

(iii) Spätestens mit der Veröffentlichung der Ergebnisse der ersten PISA-Studie ist zudem klar geworden, dass vor allem die soziale Herkunft (Migrationshintergrund, innerfamiliäres Klima, Geschlecht, Bildungshintergrund der Eltern) über den Schulerfolg entscheidet. Dementsprechend ist nicht einmal auszuschließen, dass die soziale Herkunft auch über das Fernsehverhalten entscheidet und somit die Mathematikleistung nicht *wirklich* in Beziehung zum Fernsehverhalten steht, sondern lediglich die soziale Herkunft repräsentiert. Schließlich ist davon auszugehen, dass hohe Fernsehnutzungszeiten und passive Freizeitbeschäftigungen ein Charakteristikum für eine etwaige Bildungsferne der Eltern und eine mangelhafte Medienerziehung sind.

Dass die ermittelten Zusammenhänge auch für das Gymnasium gelten (kaum Jugendliche mit Migrationshintergrund, hoher Bildungsgrad der Eltern) ist allerdings ein Indiz für eine tatsächliche Beziehung zwischen Fernsehverhalten und Mathematikleistung. Dennoch sollte in zukünftigen empirischen Arbeiten der Versuch unternommen werden, die beschriebene Problematik stärker zu berücksichtigen.

Aus den oben angeführten Gründen liegt der größte Erkenntnisgewinn dieser Arbeit daher vor allem in den neuen und weiterentwickelten datenadäquaten Auswertungsverfahren. Im Unterschied zu vielen anderen mathematikdidaktischen Studien liegt eine große Stärke der vorgestellten Untersuchung nämlich in der ausführlichen methodischen Reflexion. Der Verfasser ist der Meinung, dass in einer empirischen Forschungsarbeit, Datenerhebung einerseits und Datenauswertung andererseits harmonisch zueinander im Gleichgewicht stehen müssen. Viele empirisch arbeitende Wissenschaftler stellen zwar tiefsinnige Überlegungen über die Auswahl der Fragen (Testitems) an, lassen sich aber anschließend die Struktur der beobachteten Daten von der von ihnen benutzten Auswertungsmethode vorschreiben. Eine wirkliche, wissenschaftlich haltbare Begründung der benutzten Methoden fehlt hier meist. Sie besteht oft lediglich in dem Hinweis auf die Üblichkeit der Methoden, die sich in ähnlichen Situationen immer wieder bewährt haben. Natürlich können sie sich nicht nicht bewähren, da sie sich selber rechtfertigen, da immer etwas Interpretierbares herauskommt. Die Methode sagt dann, dass das, was herausgekommen ist, vernünftig ist. Sie schreibt vor, was vernünftig ist. Der empirisch arbeitende Wissenschaftler gibt so die Kontrolle über die Daten vollständig auf.

Daher wurde in dieser Dissertation jede Auswertungsmethode in aller Strenge gerechtfertigt. Zu diesem Zweck wurden zunächst der mögliche Informationsverlust und die mangelnde

Adäquatheit faktorenanalytischer Methoden diskutiert. Da in der Literatur und den bekannten Programmpakten (Excel, SPSS, R) keine *wirklich* adäquaten Methoden zur Verfügung stehen, die dem Niveau ordinalskalierter Daten angepasst sind, wurden bestehende Methoden weiterentwickelt und programmiert. Auf Basis eigener Beweise verschiedener mathematischer Sätze konnte so u. a. nachgewiesen werden, dass durch die Dichotomisierung der Variablen mit Hilfe der begriffsanalytischen Messtheorie nach WILLE der Informationsverlust praktisch verschwindet und die Daten nicht weiter in unverantwortlicher Weise verfälscht werden. Um die Funktionalität dieser Methode zu überprüfen, wurde zudem ein an die formale Begriffsanalyse angelehntes Verfahren entwickelt. Darüber hinaus ist es gelungen, ein adäquates Verfahren zur Beseitigung auftretender *missing values* herzuleiten, bei welchem der höchstmögliche Verlust, der entstehen kann, wenn man die durch den *missing value* entstandene Lücke in der Datenmatrix durch das Einfügen eines *falschen* Wertes schließt, minimiert wird.

Die empirische Studie diente im Zuge dessen nicht nur zur Erforschung möglicher Interdependenzen zwischen Fernsehverhalten und Mathematikleistung. Aufgrund der eigenen empirischen Untersuchung war es zudem erstmals möglich, die entwickelten Verfahren auf Basis eines großen und realen Datensatzes im Hinblick auf die praktische Anwendbarkeit zu testen. Die daraus resultierenden Ergebnisse deuten darauf hin, dass sich die methodischen Weiterentwicklungen in der Praxis tatsächlich bewährt haben, was durch folgende drei Aspekte bekräftigt werden soll:

(i) Im Zuge der Extraktion der latenten Faktoren des Fernsehkonsums stellte sich heraus, dass Verfahren der Faktorenanalyse bei ordinal skalierten Daten tatsächlich lediglich auf Basis dichotomer Zufallsvariablen sinnvolle Ergebnisse liefern. Wird eine übliche Faktorenanalyse durchgeführt, werden statt den fünf Faktoren zehn inhaltlich kaum zu interpretierende Faktoren ermittelt, so dass das Antwortverhalten der Probanden bei Anwendung gängiger Verfahren wahrhaftig falsch interpretiert wird. Möglichen Kritikern der provokanten Beispiele aus dem mathematischen Kapitel kann somit gezeigt werden, dass die modifizierte Faktorenanalyse auch angewendet auf reale und weniger extreme Datensätze erfolgreiche Arbeit leistet, wohingegen die Standardverfahren nachweislich versagen würden.

(ii) Durch die Erweiterung des Hauptsatzes der formalen Begriffsanalyse auf beliebige Ordnungen konnten die durch die weiterentwickelte Faktorenanalyse gewonnenen Ergebnisse zusätzlich bestätigt werden. Auch dieses Verfahren konnte erstmals an einem großen und realen Datensatz getestet werden und lieferte sowohl methodisch als auch inhaltlich äußerst zufriedenstellende Ergebnisse.

(iii) Das selbst entwickelte und an ein bekanntes Optimierungsproblem aus der mathematischen Spieltheorie angelehnte Verfahren zur Schätzung der *missing va-*

lues konnte ebenfalls angewendet auf einen realen Datensatz überzeugen. Die *missing values* wurden so adäquat geschätzt, dass selbst ein Datensatz mit vielen fehlenden Werten sinnvoll interpretiert werden konnte.

Aus diesen Gründen soll diese Dissertation mit einer Aussage abgeschlossen werden, die einem berühmten Zitat, welches häufig Winston Churchill zugeschrieben wird, angelehnt ist:

„Traue keiner Statistik, die nicht auf Basis datenadäquater Methoden erhoben wurde!"

Anhang

Beweise

Beweis zu Satz 4.1:

(i) $\Rightarrow$ (ii): - Trivial -

(ii) $\Rightarrow$ (iii): Angenommen, $\mathbb{W}$ enthält mehr als zwei Elemente. Dann existieren reelle Zahlen $x \in \mathbb{W}$, $y \in \mathbb{W}$ und $z \in \mathbb{W}$ mit $x < y < z$. Darüber hinaus gibt es wegen Postulat **P1** paarweise disjunkte Mengen A, B und C, deren Wahrscheinlichkeiten größer als Null sind und deren Vereinigung mit der Grundmenge $\boldsymbol{\Omega}$ übereinstimmt. Postulat **P2** garantiert nun die Existenz von Zufallsvariablen X und Y, die für alle $\omega \in \Omega$ durch

$$X(\omega) := \begin{cases} x, & \text{falls } \omega \in C \\ y, & \text{falls } \omega \in B \\ z, & \text{falls } \omega \in A \end{cases}$$

und

$$Y(\omega) := \begin{cases} y, & \text{falls } \omega \in C \\ x, & \text{falls } \omega \in B \cup A \end{cases}$$

definiert sind. Wegen Aussage (ii) dürfen wir o.E. annehmen, dass $x = 0$, $y = 1$ und $z = 2$ gilt. Dann folgt:

$$cor(X,Y) = \frac{E(X \cdot Y) - E(X) \cdot E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}} =$$

$$\frac{\left(P(B) + 2P(A)\right) \cdot P(C)}{\sqrt{P(B) + 4P(A) - P(B)^2 - 4P(B)P(A) - 4P(A)^2} \cdot \sqrt{P(C) - P(C)^2}}.$$

Unser Ziel ist es nun eine reelle Zahl $k > 2$ so zu finden, dass für einen Ordnungsautomorphismus $T:(\mathbb{R},\leq) \mapsto (\mathbb{R},\leq)$ mit $T(0) = 0$, $T(1) = 1$ und $T(2) = k$ die Korrelationskoeffizienten $cor(X,Y)$ und $cor(T \circ X, T \circ Y) = cor(T \circ X, Y)$ verschieden sind. Damit hätten wir dann den gewünschten Widerspruch erhalten und die Implikation (ii) $\Rightarrow$ (iii) bewiesen. Für den gesuchten Ordnungsautomorphismus $T:(\mathbb{R},\leq) \mapsto (\mathbb{R},\leq)$ soll also gelten

$$cor\left(X,Y\right)=\frac{\left(P\left(B\right)+2P\left(A\right)\right)\cdot P\left(C\right)}{\sqrt{P\left(B\right)+4P\left(A\right)-P\left(B\right)^2-4P\left(B\right)P\left(A\right)-4P\left(A\right)^2}\cdot\sqrt{P\left(C\right)-P\left(C\right)^2}}\neq$$

$$cor\left(T\circ X,Y\right)=\frac{\left(P\left(B\right)+2P\left(A\right)\right)\cdot P\left(C\right)}{\sqrt{P\left(B\right)+k^2P\left(A\right)-P\left(B\right)^2-2kP\left(B\right)P\left(A\right)-k^2P\left(A\right)^2}\cdot\sqrt{P\left(C\right)-P\left(C\right)^2}},$$

was zu der Ungleichung

$$\frac{P\left(B\right)^2+4P\left(B\right)P\left(A\right)+4P\left(A\right)^2}{P\left(B\right)+4P\left(A\right)-P\left(B\right)^2-4P\left(B\right)P\left(A\right)-4P\left(A\right)^2}\neq$$
$$\frac{P\left(B\right)^2+2kP\left(B\right)P\left(A\right)+k^2P\left(A\right)^2}{P\left(B\right)+k^2P\left(A\right)-P\left(B\right)^2-2kP\left(B\right)P\left(A\right)-k^2P\left(A\right)^2}$$

oder zu der Ungleichung

$$\frac{P\left(B\right)^2+4P\left(B\right)P\left(A\right)+4P\left(A\right)^2}{P\left(B\right)^2+2kP\left(B\right)P\left(A\right)+k^2P\left(A\right)^2}\neq\frac{P\left(B\right)+4P\left(A\right)-P\left(B\right)^2-4P\left(B\right)P\left(A\right)-4P\left(A\right)^2}{P\left(B\right)+k^2P\left(A\right)-P\left(B\right)^2-2kP\left(B\right)P\left(A\right)-k^2P\left(A\right)^2}$$

äquivalent ist.

Wir setzen nun $s:=P\left(B\right)^2+4P\left(B\right)P\left(A\right)+4P\left(A\right)^2$, $p:=P\left(B\right)^2$, $t:=P\left(B\right)+4P\left(A\right)-P\left(B\right)^2$ $-4P\left(B\right)P\left(A\right)-4P\left(A\right)^2$, $q:=P\left(B\right)-P\left(B\right)^2$ und darüber hinaus $a:=P\left(A\right)$ und $b:=P\left(B\right)$ und erhalten auf diese Weise die Ungleichung

$$\frac{s}{p+2kba+k^2a^2}\neq\frac{t}{q+k^2a-2kba-k^2a^2}.$$

Jetzt nehmen wir - im Gegenteil - an, dass für alle reelle Zahlen $k>2$ stets

$$\frac{s}{p+2kba+k^2a^2}=\frac{t}{q+k^2a-2kba-k^2a^2}$$

gilt. Dann folgt für alle reelle Zahlen $k>2$: $sq-tp=ka\left(2ba+sb\right)+k^2a\left(t-s+sa\right)$.

Nun unterscheiden wir zwei Fälle:

1. Fall: $t - s + sa < 0$. In diesem Fall ist der Ausdruck $ka\left(2ba + sb\right) + k^2a\left(t - s + sa\right)$ mit wachsendem k für alle reellen Zahlen $k > 2$ mit $k\left(s - sa - t\right) > 2ba + sb$ streng monoton fallend.

2. Fall: $t - s + sa \geq 0$. In diesem Fall ist der Ausdruck $ka\left(2ba + sb\right) + k^2a\left(t - s + sa\right)$ mit wachsendem k für alle reellen Zahlen $k > 2$ mit $k\left(s - sa - t\right) > 2ba + sb$ streng monoton steigend.

Beide Fälle implizieren, dass nicht für alle reelle Zahlen > 2 die Gleichung

$$\frac{s}{p + 2kba + k^2a^2} = \frac{t}{q + k^2a - 2kba - k^2a^2}$$

gelten kann, was den Beweis der Implikation (ii) $\Rightarrow$ (iii) beschließt.

(iii) $\Rightarrow$ (i): Wir dürfen o.E. annehmen, dass $\mathbb{W}$ genau zwei Elemente $x < y$ enthält. Dann betrachten wir beliebige Zufallsvariablen X, Y, deren Werte in $\mathbb{W}$ liegen und wählen irgendwelche voneinander verschiedenen reelle Zahlen v und w. Sei nun $B : \mathbb{R} \mapsto \mathbb{R}$ eine Bijektion mit $B\left(x\right) = v$ und $B\left(Y\right) = w$. Wir müssen zeigen, dass $cor\left(X,Y\right) = cor\left(B \circ X, B \circ Y\right)$ gilt. Dazu rechnen wir einfach nach, dass $B \circ X = \frac{v - w}{x - y} \cdot X + v - x \cdot \frac{v - w}{x - y}$ und $B \circ Y = \frac{v - w}{x - y} \cdot Y + v - x \cdot \frac{v - w}{x - y}$ gilt (natürlich weiß bereits jeder Schüler, dass unter allen bijektiven Abbildungen $B : \mathbb{R} \mapsto \mathbb{R}$ mit $B\left(x\right) = v$ und $B\left(Y\right) = w$ genau eine, durch eine Geradengleichung der Form $g := m \cdot k + t$, beschreibbare Bijektion existiert). Die bekannten Invarianzeigenschaften des Korrelationskoeffizienten liefern dann die gewünschte Gleichung.

Q.E.D.

Beweis zu Satz 4.2:

ad (i): Wir gehen indirekt vor, nehmen also an, dass $\mathbb{W}$ mehr als zwei Elemente enthält und zeigen, dass „s" das Postulat **SP1** nicht erfüllen kann. In analoger Weise folgt dann, dass „s" auch dem Postulat **SP2** nicht genügen kann. Um zu beweisen, dass „s" das Postulat **SP1** nicht erfüllen kann, setzen wir voraus, dass für alle affinen Transformationen der Form $m \cdot k + t$ $\left(m > 0,\ t \text{ beliebig}\right)$ auf $\mathbb{R}$ die Gleichung $s\left(X,X\right) = s\left(m \cdot X + t, X\right)$ gilt. Der Fall, dass für alle affinen Transformationen der Form $m \cdot k + t$ $\left(m > 0,\ t \text{ beliebig}\right)$ auf $\mathbb{R}$ die Gleichung $s\left(X,X\right) = s\left(X, m \cdot X + t\right)$ gültig ist, kann in analoger Weise behandelt werden. Wir fahren fort, indem wir die gleiche Notation wie im Beweis von **Satz 4.1** benutzen und betrachten die Zufallsvariable $X \in \mathfrak{X}$, die durch

$$X(\omega) := \begin{cases} x, & \text{falls } \omega \in \mathrm{C} \cup \mathrm{B} \\ z, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$ definiert ist. Dann wählen wir die durch $T(k) := \frac{y-x}{z-x} \cdot k + y - z \cdot \frac{y-x}{z-x}$ für alle $k \in \mathbb{R}$ auf $\mathbb{R}$ definierte affine Transformation. Es gilt:

$$T \circ X(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ y, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$. Weiter gilt: $T \circ X \prec X \precsim X$. Wenn Postulat **SP1** erfüllt wäre, so müsste also folgen, dass $s(T \circ X, X) < s(X, X)$ wahr ist. Andererseits impliziert die entsprechende Voraussetzung von Aussage (i) jedoch, dass $s(T \circ X, X) = s(X, X)$ gilt. Dieser Widerspruch beweist Aussage (i) des Satzes.

ad (ii): Wie beim Beweis von Aussage (i) argumentieren wir indirekt. Dazu nehmen wir an, dass $\mathbb{W}$ mindestens drei Elemente $x < y < z$ enthält. Darüber hinaus unterstellen wir, dass „s" symmetrisch und permutationsunabhängig ist. Unter diesen Voraussetzungen zeigen wir anschließend durch Kontraposition, dass „s" das Postulat **SP1** nicht erfüllen kann. Eine vollständig analoge Argumentation erlaubt es uns dann zu folgern, dass für „s" auch das Postulat **SP2** nicht erfüllbar ist. Wir nehmen also - im Gegenteil - an, dass „s" das Postulat **SP1** erfüllt und zeigen in einem ersten Schritt, dass „s" für alle Zufallsvariablen $X \in \mathfrak{X}$, $Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$, die höchstens die Werte x, y und z annehmen, auch dem Postulat **SP2** genügt. Seien dazu $X \precsim Y \prec Z$ entsprechende Zufallsvariable von $\mathfrak{X}$. Wir müssen die Ungleichung $s(X, Z) < s(X, Y)$ verifizieren. Hierzu betrachten wir eine bijektive Abbildung $B : \mathbb{W} \mapsto \mathbb{W}$ mit $B(x) = z$, $B(y) = y$ und $B(z) = x$. Es folgt, dass $B \circ Z \prec B \circ Y \precsim B \circ X$ gilt. Da „s" das Postulat **SP1** erfüllt, können wir daraus schließen, dass $s(B \circ Z, B \circ X) < s(B \circ Y, B \circ X)$ gilt. Die Symmetrie von „s" impliziert somit die Gültigkeit der Ungleichung $s(B \circ X, B \circ Z) < s(B \circ X, B \circ Y)$. Aus der Permutationsunabhängigkeit von „s" folgt dann die gewünschte Ungleichung $s(X, Z) < s(X, Y)$. Im zweiten Schritt übernehmen wir die Notation des Beweises von **Satz 4.1**, um Zufallsvariable $X \in \mathfrak{X}$, $Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ durch

$$X(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ z, & \text{falls } \omega \in A \end{cases},$$

$$Y(\omega) := \begin{cases} x, & \text{falls } \omega \in C \\ y, & \text{falls } \omega \in B \\ z, & \text{falls } \omega \in A \end{cases}$$

und

$$Z(\omega) := \begin{cases} x, & \text{falls } \omega \in C \\ z, & \text{falls } \omega \in B \cup A \end{cases}$$

für alle $\omega \in \Omega$ zu definieren. Es folgt, dass $X \prec Y \prec Z$ gilt. Da wir vorausgesetzt haben, dass „s" das Postulat **SP1** erfüllt, können wir aus diesen Ungleichungen schließen, dass $s(X,Z) < s(Y,Z)$ gelten muss. Nun wählen wir eine bijektive Abbildung $B' : \mathbb{W} \mapsto \mathbb{W}$ mit $B'(x) = y$, $B'(y) = z$ und $B'(z) = x$ aus. Dann folgt:

$$B' \circ X(\omega) := \begin{cases} y, & \text{falls } \omega \in C \cup B \\ x, & \text{falls } \omega \in A \end{cases},$$

$$B' \circ Y(\omega) := \begin{cases} y, & \text{falls } \omega \in C \\ z, & \text{falls } \omega \in B \\ x, & \text{falls } \omega \in A \end{cases}$$

und

$$B' \circ Z(\omega) := \begin{cases} y, & \text{falls } \omega \in C \\ x, & \text{falls } \omega \in B \cup A \end{cases}$$

für alle $\omega \in \Omega$. Daraus resultieren die Ungleichungen $B' \circ Z \prec B' \circ X \prec B' \circ Y$. Da „s" wie im ersten Schritt gezeigt wurde, für diese Zufallsvariablen auch das Postulat **SP2** erfüllt, können wir folgern, dass die Ungleichung $s(B' \circ Z, B' \circ Y) < s(B' \circ Z, B' \circ X)$ gültig ist. Die Symmetrie von „s" impliziert daher die Richtigkeit der Ungleichung $s(B' \circ Y, B' \circ Z) < s(B' \circ X, B' \circ Z)$. Andererseits gilt jedoch $s(X,Z) < s(Y,Z)$, woraus wegen der Permutationsunabhängigkeit von „s" abgeleitet werden darf, dass $s(B' \circ X, B' \circ Z) < s(B' \circ Y, B' \circ Z)$ gelten muss. Der auf diese Weise erhaltene Widerspruch beendet den Beweis von Aussage (ii).

ad (iii): Wir gehen wiederum indirekt vor, nehmen als an, dass $\mathbb{W}$ mindestens vier Elemente $v < x < y < z$ enthält. Dann nehmen wir an, dass „s" das Postulat **SP1** erfüllt und zeigen, dass „s" unter diesen Voraussetzungen der in Aussage (ii) geforderten Invarianzeigenschaft nicht genügen kann. In analoger Weise gehen wir vor, wenn „s" das Postulat **SP2** erfüllt. Wir benutzen die Notation des Beweises von **Satz 4.1**, um Zufallsvariable $Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ durch

$$Y(\omega) := \begin{cases} v, & \text{falls } \omega \in C \cup B \\ y, & \text{falls } \omega \in A \end{cases}$$

und

$$Z(\omega) := \begin{cases} v, & \text{falls } \omega \in C \cup B \\ z, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$ zu definieren. Nun betrachten wir eine stückweise lineare und streng monoton steigende Transformation $T : (\mathbb{R}, \leq) \mapsto (\mathbb{R}, \leq)$ mit $T(v) = v$, $T(x) = x$ und $T(z) = z$. Dann gilt: $T \circ Y \prec Y \prec Z = T \circ Z$. Mit Hilfe des Postulats **SP1** folgt daraus die Ungleichung $s(T \circ Y, Z) < s(Y, Z)$. Wenn also $s(Y, Z) = s(T \circ Y, T \circ Z)$ gelten würde, so würde auch folgen, dass $s(T \circ Y, Z) = s(T \circ Y, T \circ Z) = s(Y, Z)$ gilt, was der Ungleichung $s(T \circ Y, Z) < s(Y, Z)$ widerspricht. Damit ist Aussage (iii) bewiesen.

ad (iv): Wir nehmen - im Gegenteil - an, dass $\mathbb{W}$ fünf äquidistante Punkte $v < w < x < y < z$ enthält und beweisen, dass „s" unter der Voraussetzung der in Aussage (iv) unterstellten Invarianzeigenschaften nicht beide Postulate **SP1** und **SP2** erfüllen kann. Dazu benutzen wir die Notation des Beweises von **Satz 4.1** und definieren zwei Zufallsvariable $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ durch

$$X(\omega) := \begin{cases} w, & \text{falls } \omega \in C \cup B \\ x, & \text{falls } \omega \in A \end{cases}$$

und

$$Y(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ y, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$. Dann wählen wir die durch $T(k) := \frac{x - v}{x - w} \cdot k + v - w \cdot \frac{x - v}{x - w}$ für alle $k \in \mathbb{R}$ definierte affine Transformation auf $\mathbb{R}$. Es gilt

$$T \circ X(\omega) := \begin{cases} v, & \text{falls } \omega \in C \cup B \\ x, & \text{falls } \omega \in A \end{cases}$$

und

$$T \circ Y(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ z, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$. Daraus folgen die Beziehungen $T \circ X \prec X \prec Y \prec T \circ Y$. Die Postulate **SP1** und **SP2** implizieren nun die Ungleichungen $s(T \circ X, T \circ Y) \underset{SP1}{<} s(X, T \circ Y) \underset{SP2}{<} s(X,Y)$. Andererseits wird jedoch vorausgesetzt, dass $s(X,Y) = s(T \circ X, T \circ Y)$ gilt. Der so erhaltene Widerspruch beweist Aussage (iv) des Satzes.

Q.E.D.

Beweis zu Satz 4.3:

Der Beweis des Satzes ist wie folgt organisiert: Zunächst beweisen wir die Kette (iii) $\Rightarrow$ (i) $\Rightarrow$ (vii) $\Leftrightarrow$ (vi) von Implikationen und Äquivalenzen. Danach verifizieren wir die Äquivalenz (vii) $\Leftrightarrow$ (v), um schließlich noch die Implikation (vii) $\Rightarrow$ (iii) zu zeigen. Die dann noch zu beweisende Kette (iv) $\Rightarrow$ (ii) $\Rightarrow$ (vii) $\Rightarrow$ (iv) von Implikationen folgt in analoger Weise wie die Implikationskette (iii) $\Rightarrow$ (i) $\Rightarrow$ (vii) $\Rightarrow$ (iii). Ihr Nachweis kann daher vernachlässigt werden.

(iii) $\Rightarrow$ (i): - Trivial -

(i) $\Rightarrow$ (vii): Wir argumentieren wiederum indirekt, nehmen also - im Gegenteil - an, dass $\mathbb{W}$ mindestens drei (verschiedene) Elemente $x < y < z$ enthält. Mit Hilfe der im Beweis der Implikation (ii) $\Rightarrow$ (iii) von **Satz 4.1** verwendeten Notation und Schlussweise lassen sich dann für alle $\omega \in \Omega$ Zufallsvariable $X \in \mathfrak{X}$, $Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ durch

$$X(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ y, & \text{falls } \omega \in A \end{cases},$$

$$Y(\omega) := \begin{cases} x, & \text{falls } \omega \in C \\ y, & \text{falls } \omega \in B \\ z, & \text{falls } \omega \in A \end{cases}$$

und

$$Z(\omega) := \begin{cases} y, & \text{falls } \omega \in C \cup B \\ z, & \text{falls } \omega \in A \end{cases}$$

definieren. Man verifiziert sofort, dass die Voraussetzungen von Postulat **WSP1**, nämlich $X \prec Y \precsim Z$, erfüllt sind. Da keine reellen Zahlen $m > 0$, t beliebig mit $Z = m \cdot Y + t$ fast sicher existieren, gilt die Ungleichung $cor(Y, Z) < 1$. Die im Beweis der Implikation (iii) $\Rightarrow$ (i) von **Satz 4.1** benutzte Argumentation beweist jedoch, dass $cor(X, Z)$ mit 1 übereinstimmt. Damit haben wir den gewünschten Widerspruch erhalten und die Implikation bewiesen.

(vii) $\Rightarrow$ (vi): Wegen **Satz 4.1** dürfen wir o.E. annehmen, dass $\mathbb{W} = \{0,1\}$ gilt. Dann folgt:

$cor(X,Y) = 1$

$\Leftrightarrow$ es existieren eine positive reelle Zahl m und eine beliebige reelle Zahl t mit $X = m \cdot Y + t$ fast sicher

$\Leftrightarrow$ (da $\mathbb{W} = \{0,1\}$ und m positiv) $1 = m \cdot 1 + t$ und $0 = m \cdot 0 + t$ fast sicher

$\Leftrightarrow$ $m = 1$ und $t = 0$ fast sicher

$\Leftrightarrow$ $P(\{X = Y\}) = 1$.

(vi) $\Rightarrow$ (vii): Wir schließen direkt. Angenommen, $\mathbb{W}$ enthält mindestens drei Elemente $x < y < z$. Dann benutzen wir die Notation und Argumentation des Beweises der Implikation (ii) $\Rightarrow$ (iii) von **Satz 4.1**, um Zufallsvariable $X \in \mathfrak{X}$ und $Y \in \mathfrak{X}$ durch

$$X(\omega) := \begin{cases} x, & \text{falls } \omega \in C \cup B \\ y, & \text{falls } \omega \in A \end{cases}$$

und

$$Y(\omega) := \begin{cases} y, & \text{falls } \omega \in C \cup B \\ z, & \text{falls } \omega \in A \end{cases}$$

für alle $\omega \in \Omega$ zu definieren. Mit Hilfe des im Beweis der Implikation (iii) $\Rightarrow$ (i) von **Satz 4.1** verwendeten Arguments erhalten wir die Gleichung $cor(X, Y) = 1$. Wegen $P(\{X = Y\}) = 0$, haben wir so den gewünschten Widerspruch hergeleitet.

(vii) $\Rightarrow$ (v): Es genügt, die Gültigkeit von Postulat **PA*** zu beweisen. Da es sich bei den Zufallsvariablen X, Y und Z um dichotome Zufallsvariable handelt, können wir wegen **Satz**

4.1 beliebige Mengen $C \subset A$ und $B \subset A$ mit $P(C) < P(B)$ betrachten und die Zufallsvariablen X, Y und Z mit den entsprechenden Indikatorvariablen $i_C \precsim i_A$ und $i_B \precsim i_A$ identifizieren. Weil die Zufallsvariablen X und Z nicht fast sicher konstant sind, folgt $0 < P(C) < P(A) < 1$. Die Ungleichung $cor(X,Z) < cor(Y,Z)$ ist dann zur Ungleichung

$$\frac{E(i_C \cdot i_A) - E(i_C) \cdot E(i_A)}{\sqrt{E(i_C^{\,2}) - E(i_C)^2} \cdot \sqrt{E(i_A^{\,2}) - E(i_A)^2}} < \frac{E(i_B \cdot i_A) - E(i_B) \cdot E(i_A)}{\sqrt{E(i_B^{\,2}) - E(i_B)^2} \cdot \sqrt{E(i_A^{\,2}) - E(i_A)^2}}$$

bzw. der Ungleichung

$$\frac{P(C) - P(C) \cdot P(A)}{\sqrt{P(C) - P(C)^2} \cdot \sqrt{P(A) - P(A)^2}} < \frac{P(B) - P(B) \cdot P(A)}{\sqrt{P(B) - P(B)^2} \cdot \sqrt{P(A) - P(A)^2}}$$

äquivalent.

Nun gilt:

$$\frac{P(C) - P(C) \cdot P(A)}{\sqrt{P(C) - P(C)^2} \cdot \sqrt{P(A) - P(A)^2}} < \frac{P(B) - P(B) \cdot P(A)}{\sqrt{P(B) - P(B)^2} \cdot \sqrt{P(A) - P(A)^2}}$$

$$\Leftrightarrow \quad \frac{P(C)(1 - P(A))}{\sqrt{P(C)(1 - P(C))}} < \frac{P(B)(1 - P(A))}{\sqrt{P(B)(1 - P(B))}}$$

$$\Leftrightarrow \quad \frac{P(C)^2}{P(C)(1 - P(C))} < \frac{P(B)^2}{P(B)(1 - P(B))}$$

$$\Leftrightarrow \quad \frac{P(C)}{1 - P(C)} < \frac{P(B)}{1 - P(B)}$$

$$\Leftrightarrow \quad P(C) - P(B) \cdot P(C) < P(B) - P(B) \cdot P(C)$$

$\Leftrightarrow \quad P(C) < P(B)$, was die Gültigkeit von Postulat **PA*** beweist.

(v) $\Rightarrow$ (vii): - Trivial -

(vii) $\Rightarrow$ (iii): Diese Implikation folgt mit den gleichen Argumenten, wie sie bereits im Beweis der Implikation (vii) $\Rightarrow$ (v) benutzt wurden.

Q.E.D.

Beweis zu Satz 4.4:
(i) $\Rightarrow$ (ii): Der Beweis dieser Implikation basiert im Wesentlichen auf folgendem Lemma, das als Proposition 5.1 in Herden und Herden (*preprint* 2013) bewiesen wurde.

Lemma 1:
Sei $(\Omega, \mathcal{A}, \mu)$ *ein endlicher Maßraum. Dann existieren eine nichtnegative reelle Zahl* $c \leq \mu(\Omega)$ *und eine Reihe* $\sum_{k=0}^{\infty} a_k$ *nichtnegativer reeller Zahlen* a_k*, die gegen* $\mu(\Omega) - c$ *konvergiert, so dass folgende Aussagen erfüllt sind:*

(i) *Sei* $add: [0,c] \times \left\{ y \in \mathbb{R}^{\geq 0} \mid \exists T \subset \mathbb{N} \left(y = \sum_{k \in T} a_k \right) \right\} \mapsto \mu(\mathcal{A})$ *die durch* $add(x,y) := x + y$ *für alle Paare* $(x,y) \in [0,c] \times \left\{ y \in \mathbb{R}^{\geq 0} \mid \exists T \subset \mathbb{N} \left(y = \sum_{k \in T} a_k \right) \right\}$ *definierte stetige Funktion. Dann gilt* $\mu(\mathcal{A}) = add\left([0,c] \times \left\{ y \in \mathbb{R}^{\geq 0} \mid \exists T \subset \mathbb{N} \left(y = \sum_{k \in T} a_k \right) \right\} \right) = \left\{ x + y \mid x \in [0,c] \wedge \exists T \subset \mathbb{N} \left(y = \sum_{k \in T} a_k \right) \right\}$.

(ii) $\left\{ y \in \mathbb{R}^{\geq 0} \mid \exists T \subset \mathbb{N} \left(y = \sum_{k \in T} a_k \right) \right\}$ *und* $\mu(\mathcal{A})$ *sind kompakte Teilmengen reeller Zahlen.*

Zum Beweis der Implikation (i) $\Rightarrow$ (ii) von **Satz 4.4** benötigen wir über Lemma 1 hinaus noch folgende Definition und folgendes Lemma.

Definition:
Eine Abbildung $s : \mathfrak{X} \times \mathfrak{X} \mapsto [-1,1]$ heißt *I-stetig* (der Buchstabe „I" soll an Indikator erinnern), wenn für alle Mengen $E \in \mathcal{A}$ und alle monoton wachsenden Folgen $D_0 \subset D_1 \subset D_2 \subset \ldots \subset D_n \subset D_{n+1} \subset \ldots \subset \ldots$ von Teilmengen von E, deren Wahrscheinlichkeiten größer als Null sind, die Gleichheit von $s\left(i_{\bigcup_{n \in \mathbb{N}} D_n}, i_E \right)$ und $\lim_{n \to \infty} s\left(i_{D_n}, i_E \right)$ folgt.

Lemma 2:
$cor : \mathfrak{X} \times \mathfrak{X} \mapsto [-1,1]$ *ist I-stetig.*

Beweis:
Die Aussage des Lemmas folgt unmittelbar aus der Definition des Korrelationskoeffizienten mit Hilfe der Limitenregeln für Folgen von Quotienten, für die die entsprechenden Limitenregeln implizieren, dass sowohl die Folge der Zähler als auch die Folge der Nenner konvergiert.

Natürlich darf die Folge der Nenner dabei nicht gegen 0 konvergieren.

Q.E.D.

Wir haben nun alle Vorbereitungen getroffen, um den Beweis der Implikation (i) $\Rightarrow$ (ii) von **Satz 4.4** beenden zu können. Zunächst einmal impliziert Aussage (ii) von **Satz 4.2** die Gültigkeit von Bedingung **W**. Da Bedingung **B1** die Permutationsunabhängigkeit von „s" bezüglich dichotomer Zufallsvariablen beschreibt, muss zum Beweis der Implikation (i) $\Rightarrow$ (ii) von **Satz 4.4** nur noch Bedingung **B2** verifiziert werden. Dazu benutzen wir den Beweis der Implikation (vii) $\Rightarrow$ (v) von **Satz 4.3**. Seien also beliebige Zufallsvariable $X \in \mathfrak{X}$, $Y \in \mathfrak{X}$ und $Z \in \mathfrak{X}$ mit $X \precsim Z$ und $Y \precsim Z$ vorgegeben. Dann implizieren die im Beweis der Implikation (vii) $\Rightarrow$ (v) von **Satz 4.3** durchgeführten Äquivalenzumformungen mit Hilfe von Bedingung **W**, der daraus resultierenden Skalierungsunabhängigkeit des Korrelationskoeffizienten und den Bedingungen **PA1** und **PA2**, dass die Äquivalenzen $s\left(X,Z\right) \le s\left(Y,Z\right) \Leftrightarrow P\left(\left\{X = Z\right\}\right) \le P\left(\left\{Y = Z\right\}\right) \Leftrightarrow cor\left(X,Z\right) \le cor\left(Y,Z\right)$ gelten. Wenn wir daher für eine fest vorgegebene Zufallsvariable $Z \in \mathfrak{X}$ die Menge $I_Z := \left\{cor\left(X,Z\right) \mid X \in \mathfrak{X} \text{ und } X \precsim Z\right\}$ betrachten, so garantieren diese Äquivalenzen die Existenz einer ordnungserhaltenden Abbildung $t_Z : \left(I_Z, \le\right) \mapsto \left(\left[-1,1\right], \le\right)$ mit $s\left(X,Z\right) = t_Z\left(cor\left(X,Z\right)\right)$ für alle Zufallsvariablen $X \in \mathfrak{X}$ mit $X \precsim Z$. Der Beweis der Implikation (i) $\Rightarrow$ (ii) von **Satz 4.4** ist also abgeschlossen, wenn wir noch nachweisen können, dass sich t_Z zu einer ordnungserhaltenden Abbildung $T_Z : \left(\left[-1,1\right], \le\right) \mapsto \left(\left[-1,1\right], \le\right)$ fortsetzen lässt. Dazu genügt es zu zeigen, dass I_Z eine abgeschlossene Teilmenge von $\left[-1,1\right]$ ist. In diesem Falle lässt sich t_Z nämlich durch stückweise lineare (streng monoton steigende) Funktionen zu der gewünschten Abbildung $T_Z : \left(\left[-1,1\right], \le\right) \mapsto \left(\left[-1,1\right], \le\right)$ erweitern. Sei daher z ein Randpunkt von I_Z. Um die Beziehung $z \in I_Z$ zu verifizieren, vergewissern wir uns der Existenz einer Folge $\left(X_n\right)_{n\in\mathbb{N}}$ von Zufallsvariablen $X_n \in \mathfrak{X}$ und $X_n \precsim Z$ mit $\lim\limits_{n\to\infty} cor\left(X_n, Z\right) = z$. Wegen der Bedingung **W**, die die Skalierungsunabhängigkeit des Korrelationskoeffizienten garantiert, dürfen wir die Folge $\left(X_n\right)_{n\in\mathbb{N}}$ von Zufallsvariablen mit einer Folge $\left(i_{V_n}\right)_{n\in\mathbb{N}}$ von Indikatorvariablen identifizieren. Dies bedeutet auch, dass Z mit einer Indikatorvariablen i_W identifiziert werden kann. Für die entsprechenden Mengen $V_n \in \mathcal{A}$ und $W \in \mathcal{A}$ kann o.E. $P\left(V_n\right) < P\left(V_{n+1}\right) < P\left(W\right)$ für alle $n \in \mathbb{N}$ oder $P\left(V_n\right) > P\left(V_{n+1}\right) > P\left(W\right)$ für alle $n \in \mathbb{N}$ angenommen werden. Da $\mathcal{A}$ gegenüber Komplementbildung abgeschlossen ist, dürfen wir o.E. voraussetzen, dass für alle $n \in \mathbb{N}$ die Ungleichungen $P\left(V_n\right) < P\left(V_{n+1}\right) < P\left(W\right)$ gültig sind. Lemma 1 garantiert nun die Existenz einer Menge $H \in \mathcal{A}$, für die $P\left(H\right) = \lim\limits_{n\to\infty} P\left(V_n\right)$ gilt. Die für alle $n \in \mathbb{N}$ gültigen Ungleichungen $P\left(V_n\right) < P\left(V_{n+1}\right) < P\left(W\right)$ erlauben es uns darüber hinaus anzunehmen, dass zu jeder positiven reellen Zahl ε ein $A \in \mathcal{A}$ existiert, welches ganz in H enthalten ist und die Ungleichung $0 < P(A) < \varepsilon$ erfüllt. Diese Beobachtung impliziert mit Hilfe von Lemma 1 die Existenz einer Folge $(H_n)_{n\in\mathbb{N}}$ von Mengen $H_n \subset H$, die alle zu $\mathcal{A}$ gehören und die Beziehungen $H = \bigcup_{n\in\mathbb{N}} H_n$, $H_n \underset{\neq}{\subset} H_{n+1}$ und $P(H_n) < P(H_{n+1})$ für alle $n \in \mathbb{N}$ erfüllen. Daraus folgt mit Hilfe von Lemma 2, dass die Gleichung $cor(i_H, i_W) = \lim_{n\to\infty} cor(i_{H_n}, i_W)$ gilt. Da $cor(\cdot, i_W)$ mit den Erwartungswerten von Indikatorvariablen, die nicht größer als i_W sind, streng monoton wächst (vgl. den Beweis der Implikation (vii) $\Rightarrow$ (v) von **Satz 4.3**), folgt darüber hinaus die Gültigkeit der Gleichungen $\lim_{n\to\infty} cor(i_{H_n}, i_W) = \lim_{n\to\infty} cor(i_{V_n}, i_W) = z$. Der Punkt z ist daher in I_Z enthalten, was den Beweis der Implikation (i) $\Rightarrow$ (ii) von **Satz 4.4** abschließt.

(ii) $\Rightarrow$ (i): Wegen **Satz 4.1** und **Satz 4.3** implizieren die Bedingungen **W**, **B1** und **B2** von Aussage (i), das „s" permutationsunabhängig ist und allen in **Satz 4.3** aufgeführten Bedingungen genügt. „s" erfüllt somit insbesondere Aussage (i) des Satzes.

Q.E.D.

Beweis zu Satz 4.5:

Der Beweis von Aussage (i) ist trivial. Zum Beweis des Satzes genügt es daher zu zeigen, dass zu jeder (teilweise) geordneten Menge $(X, \leq)$ ein formaler Kontext $\mathcal{K}(G, M, I)$ existiert, so dass $(X, \leq)$ zu $(\mathcal{E}(G, M, I), \preceq)$ ordnungsisomorph ist. Sei daher eine (teilweise) geordnete Menge $(X, \leq)$ beliebig aber fest vorgegeben. Dann beweisen wir, dass $(X, \leq)$ zu $(\mathcal{E}(X, X, \leq), \preceq)$ ordnungsisomorph ist. Dazu genügt es zu zeigen, dass die Abbildung $\varphi : (X, \leq) \mapsto (\mathcal{E}(X, X, \leq), \preceq)$, die für alle $x \in X$ durch $\varphi(x) := (\{x\}^{\perp\perp}, \{x\}^{\perp})$ definiert ist, ein Ordnungsisomorphismus ist. Da die Definition von φ sofort impliziert, dass φ eine bijektive Abbildung ist, müssen wir hierzu lediglich verifizieren, dass für beliebige Element $y \in X$ und $z \in X$ die strenge Ungleichung $y < z$ die strenge Ungleichung $\{z\}^{\perp} \underset{\neq}{\subset} \{y\}^{\perp}$ impliziert, die, wegen $\{z\}^{\perp} = \{z' \in X \mid z \leq z'\} \subseteqq \{y' \in X \mid y \leq y'\} = \{y\}^{\perp}$ und $y \notin \{z\}^{\perp}$, aber trivialerweise erfüllt ist. Mit dieser Beobachtung ist die Ergänzung des Hauptsatzes der formalen Begriffsanalyse bereits bewiesen.

Q.E.D.

Beweis zu Satz 9.1:
Unter der Voraussetzung, dass sowohl $\|\cdot\|$ als auch $\|\cdot\|_C$ auf einer der in Aussage (iv) beschriebenen Menge $\mathbf{S}$ gemessen worden sind, folgt sofort das beide Distanzmaße dort ordinale Signifikanz haben. Daher genügt es zum Beweis von Satz 9.1 die Äquivalenz der Aussagen (i), (ii), (iii) und (iv) zu beweisen.
Die Äquivalenz der Aussagen (ii) und (iii) ist aus Symmetriegründen evident. Darüber hinaus folgt die Äquivalenz der Aussagen (ii) und (iv) durch eine geringfügige naheliegende Modifikation des Arguments, das zum Beweis der Äquivalenz der Aussagen (i) und (iv) benutzt werden wird. Dies meint, dass zum Beweis des Satzes lediglich die Äquivalenz der Aussagen (i) und (iv) nachgewiesen werden muss. Nun sieht man aber sofort ein, dass die Gültigkeit von Aussage (iv) hinreichend dafür ist, dass Aussage (i) gilt. Zur Beendigung des Beweises des Satzes genügt es daher zu zeigen, dass die für alle $y \in \mathbf{S}$ und alle $z \in \mathbf{S}$ angenommene Gültigkeit der Äquivalenz $\|y\| \le \|z\| \Leftrightarrow \|y\|_C \le \|z\|_C$ Aussage (iv) impliziert. Dazu zeigen wir zunächst, dass die für alle $y \in \mathbf{S}$ und alle $z \in \mathbf{S}$ postulierte Äquivalenz $\|y\| < \|z\| \Leftrightarrow \|y\|_C < \|z\|_C$ die Inklusion $\mathbf{S} \subset \bigcup_{a \in \mathbb{R}^{\ge 0}} \{0, a\}^n$ zur Folge hat. Dazu genügt es zu zeigen, dass die Gültigkeit letztgenannter Äquivalenz für alle Tupel $y \in \mathbf{S}$ und $z \in \mathbf{S}$ die Existenz eines Tupels $x = (x_1, \ldots, x_n) \in \mathbf{S}$ ausschließt, für welches Indices $1 \le i \ne j \le n$ existieren mit $x_i < x_j$. Also unterstellen wir im Gegenteil die Existenz eines solchen Tupels $x = (x_1, \ldots, x_n) \in \mathbf{S}$. Dann unterscheiden wir zwei Fälle:

1. Fall: $(\mathbb{T}, \circ)$ erfüllt für alle positiven reellen Zahlen $a < c < d < b$ und $a < g < h < b$ Bedingung **LN1**. In diesem Fall nehmen wir o.E. an, dass Indices $1 \le i \ne j \le n$ mit $0 < x_i < x_j$ so gewählt wurden, dass kein Index $1 \le k \le n$ existiert, für den $x_i < x_k < x_j$ gilt. Im nächsten Schritt setzten wir $\mathbf{I} := \{1 \le p \le n \mid x_p < x_i\}$ und $\mathbf{J} := \{1 \le q \le n \mid x_j < x_q\}$. Darüber hinaus wählen wir beliebige aber fest positive reelle Zahlen a und b, die die strengen Ungleichungen $a < x_i < x_j < b$ erfüllen. Schließlich setzen wir

$$m := \begin{cases} a, \text{ falls } \mathbf{I} = \varnothing \\ \max\{x_p \mid p \in \mathbf{I}\}, \text{ falls } \mathbf{I} \ne \varnothing \end{cases}$$

und

$$M := \begin{cases} b, \text{ falls } \mathbf{J} = \varnothing \\ \min\{x_q \mid q \in \mathbf{J}\}, \text{ falls } \mathbf{J} \ne \varnothing \end{cases}.$$

Sei $s := \left|\left\{1 \leq u \leq n \mid x_u = x_i\right\}\right|$ und $t := \left|\left\{1 \leq v \leq n \mid x_j = x_v\right\}\right|$. Dann wählen wir positive reelle Zahlen ε_1 und ε_2, so dass $s \cdot \varepsilon_1 = t \cdot \varepsilon_2$ gilt. Letztere Gleichung kürzen wir mit (*) ab. Wegen $0 < x_i < x_j$ impliziert sie die strenge Ungleichung $sx_i\varepsilon_1 < tx_j\varepsilon_2$, die mit (**) abgekürzt sei. Die eventuelle Division von ε_1 und ε_2 durch eine genügend große natürliche Zahl N erlaubt es uns o.E. zusätzlich anzunehmen, dass $m < x_i - \varepsilon_1$ und $x_j + \varepsilon_2 < M$ gilt. Letztere Annahme kürzen wir durch (***) ab. Die Ungleichungen (***) garantieren die Existenz eines Ordnungsautomorphismus $\mathrm{T} \in \mathbb{T}$, der sowohl den Gleichungen $\mathrm{T}(w) = w$ für alle nichtnegativen reellen Zahlen $w \leq m$ und $\mathrm{T}(z) = z$ für alle positiven reellen Zahlen $M \leq z$ als auch den Gleichungen $\mathrm{T}(x_i) = x_i - \varepsilon_1$ und $\mathrm{T}(x_j) = x_j + \varepsilon_2$ genügt. Jetzt setzen wir den Beweis fort, in dem wir $y := \mathrm{T}(x) = \left(\mathrm{T}(x_1), \ldots, \mathrm{T}(x_n)\right)$ setzen. Anschließend schließen wir wie folgt:

$$\|y\|^2 = \sum_{k \in \mathbf{I} \cup \mathbf{J}} x_k^2 + s\left(x_i - \varepsilon_1\right)^2 + t\left(x_j + \varepsilon_2\right)^2 = \sum_{k \in \mathbf{I} \cup \mathbf{J}} x_k^2 + s\left(x_i^2 - 2x_i\varepsilon_1 + \varepsilon_1^2\right) + t\left(x_j^2 + 2x_j\varepsilon_2 + \varepsilon_2^2\right) =$$

$$\sum_{k \in \mathbf{I} \cup \mathbf{J}} x_k^2 + sx_1^2 + tx_j^2 + 2\left(tx_j\varepsilon_2 - sx_i\varepsilon_1\right) + s\varepsilon_1^2 + t\varepsilon_2^2 = \|x\|^2 + 2\left(tx_j\varepsilon_2 - sx_i\varepsilon_1\right) + s\varepsilon_1^2 + t\varepsilon_2^2 > \|x\|^2.$$

Letztere Ungleichung folgt aus Ungleichung (**) und impliziert die strenge Ungleichung $\|x\| < \|y\|$. Mit Hilfe von Gleichung (*) und den Ungleichungen (***) folgt andererseits aber:

$$\|x\|_c = \sum_{k=1}^{n} \left|x_k\right| = \sum_{k=1}^{n} x_k = \sum_{k \in \mathbf{I} \cup \mathbf{J}} x_k + sx_i + tx_j + t\varepsilon_2 - s\varepsilon_1 = \sum_{k \in \mathbf{I} \cup \mathbf{J}} x_k + t\left(x_j + \varepsilon_2\right) + s\left(x_i - \varepsilon_1\right) =$$

$\sum_{k \in \mathbf{I} \cup \mathbf{J}} \left|x_k\right| + t\left|x_j + \varepsilon_2\right| + s\left|x_i - \varepsilon_1\right| = \|y\|_c$, was der Implikation $\|x\| < \|y\| \Rightarrow \|x\|_c < \|y\|_c$ widerspricht und den ersten Fall beendet.

2. Fall: $(\mathbb{T}, \circ)$ erfüllt für alle positiven reellen Zahlen $c < a < b < d$ und $g < a < b < h$ Bedingung **LN2**. In diesem Fall wählen wir Indices $1 \leq i \neq j \leq n$ mit $x_i = \min\left\{x_k \middle| 1 \leq k \leq n\right\}$ und $x_j = \max\left\{x_k \middle| 1 \leq k \leq n\right\}$. Dann benutzen wir Bedingung **LN2**, um das im ersten Fall verwendete Argument in entsprechend modifizierter Form auf die veränderte Situation zu übertragen. Der auf diese Weise erneut erhaltene Widerspruch beschließt dann den Beweis der Inklusion $\mathbf{S} \subset \bigcup_{a \in \mathbb{R}^{\geq 0}} \{0, a\}^n$.

Der Beweis des Satzes kann nun beendet werden, wenn wir noch zeigen können, dass keine positiven reellen Zahlen $a \neq b$ existieren, zu denen es natürliche Zahlen $1 \leq l \neq m \leq n$ und Tupel $x \in \{0, a\}_l^n$ sowie $y \in \{0, b\}_m^n$ gibt, die beide in $\mathbf{S}$ enthalten sind. O.E. gelte $l < m$. Dann vergewissern wir uns mit Hilfe von Bedingung **LN1** oder Bedingung **LN2** der Existenz eines Ordnungsautomorphismusses $\mathrm{T} \in \mathbb{T}$, für den $l \cdot \mathrm{T}(a) = m \cdot b$ gilt. Anschließend setzen wir $c := \mathrm{T}(a)$. Nun betrachten wir das Tupel $z \in \{0, c\}_l^n$, welches entsteht, wenn wir in x die positive reelle Zahl a durch c ersetzen. Die $\mathbb{T}$-Abgeschlossenheit von $\mathbf{S}$ impliziert, dass z in

S enthalten ist. Wenn wir nun zeigen können, dass $\|y\| < \|z\|$ aber $\|y\|_c = \|z\|_c$ gilt, dann war auch unsere letzte Annahme falsch, was den Beweis des Satzes beenden würde. Aus $l < m$ und $l \cdot \mathrm{T}(a) = m \cdot b$ folgt zunächst, dass $b < c$ gilt. Daraus folgt $\|y\|^2 = m \cdot b^2 < l \cdot c^2 = \|z\|^2$, was die strenge Ungleichung $\|y\| < \|z\|$ impliziert. Andererseits gilt aber $\|y\|_c = m \cdot b = l \cdot c = \|z\|_c$. Dies ist der gewünschte Widerspruch, der den Beweis des Satzes beschließt.

Q.E.D.

Fragebögen und Datensätze

Die unten aufgeführten Fragebögen und Datensätze stehen unter folgendem Link im Internet kostenlos zum Download zur Verfügung:

http://www.file-upload.net/download-8514496/Anhang.rar.html

Fragebögen

- Auswahl der tatsächlich geschauten Sendungen
- benotete Sendungen - Fernsehdauer - Freizeit (Teil I)
- Freizeitverhalten (Teil II)
- Lehrerumfrage (Teil I)
- Lehrerumfrage (Teil II)

Daten zu Kapitel 4 (Mathematische Grundlagen)

- 0-1-Daten-Converter
- Makros zur Bestimmung der missing values
- Erweiterung der formalen Begriffsanalyse

Daten zu Kapitel 6 (Datenerhebung und erste Ergebnisse)

- Datensatz Dauer und Häufigkeit des Fernsehkonsums
- Datensatz latente Faktoren des Fernsehkonsums
- Datensatz latente Faktoren des Freizeitverhaltens

Daten zu Kapitel 7 (Möglichkeiten der Leistungsmessung)

- Datensatz latente Faktoren der Mathematikleistung (exemplarisch in 0-1-Daten)
- Lehrerbewertungen

Daten zu Kapitel 8 & 9 (Die Auswertung)

- Auswertung
- Daten für die Clusteranalysen

Literaturverzeichnis

Agresti, A. (1996): An Introduction to Categorical Data Analysis, New York.

Akyün, Ö. (2012): Mögliche Zusammenhänge zwischen Fernsehkonsum einerseits und Rechenfähigkeiten andererseits bei Fünftklässlern, Essen.

Anderson, A.; Basilevsky, A; Hum, D. (1983): Measurement: Theory and techniques. In: Rossi, P.; Wright, J; Anderson, A. (Hersg.): Handbook of survey research (S. 244-251).

Backhaus, K.; Erichson, B; Plinke, W.; Weiber, R. (1996): Multivariate Analysemethoden - Eine anwendungsorientierte Einführung, Berlin/Heidelberg/New York, 8. Auflage.

Bankhofer, U. (1995): Unvollständige Daten- und Distanzmatrizen in der Multivariaten Datenanalyse, Bergisch-Gladbach.

Becke, M. (1996): Konsenskriterien und Konsistenzanforderungen für beliebige Rangordnungen und Punktbewertungen, Diss. Univ. Essen, Essen.

Becker, F.; Spörrle, M.; Försterling, F. (2003): Soziale Erwünschtheit und Skalenformat als Einflussfaktoren bei der Beantwortung von Wahrscheinlichkeitsaussagen. In: Golz, J. (Hrsg.): Experimentelle Psychologie. Abstracts der 45. Tagung experimentell arbeitender Psychologen, Lengerich, 1. Auflage.

Benninghaus, H. (2005): Einführung in die sozialwissenschaftliche Datenanalyse, Oldenbourg, 7. Auflage.

Bergs, S. (1980): Optimalität bei Clusteranalysen - Experimente zur Bewertung numerischer Klassifikationsverfahren, Münster.

Bortz, J. (1993): Statistik für Sozialwissenschaftler, Berlin, 4. Auflage.

Bortz, J.; Döring, N. (2006): Forschungsmethoden und Evaluation: Für Human und Sozialwissenschaftler, Heidelberg, 4. Auflage.

Boyd, B. (2004b): A Crow Left of the Murderer (Textheft)

Brosius, F. (1998): SPSS 8: Professionelle Statistik unter Windows, Bonn.

Brügelmann, H. (2002): Kontroversen um die Schulleistungsmessung in Deutschland. Eine fiktive Diskussion über Positionen und Perspektiven in verteilten Rollen. In: Weinert, F.E. (Hrsg.): Leistungsmessung in Schulen, S. 33-44, Weinheim, Basel und Bonn.

Bühner, M. (2006): Einführung in die Test- und Fragebogenkonstruktion (2006), München, 2. Auflage.

Cattell, R. B. (1966): The Scree Test for the Number of Factors. In: *Multivariate Behavioral Research,* 1(2), S. 245-276.

Deutsches PISA-Konsortium (2005): PSIA 2003: der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs, Münster.

Dietz, W.; Gortmaker, H. (1985): Do we fatten our children at the television set? Obesity and television viewing in children and adolescents. In: Pediatrics 75, S. 807-812.

Eells, W. C. (1977): Die Zuverlässigkeit wiederholter Benotung von aufsatzähnlichen Prüfungsarbeiten. In Ingenkamp, K. (Hrsg.): Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte, S. 167-172, Weinheim, 7. Auflage.

Ennemoser, M. (2003a): Der Einfluss des Fernsehens auf die Entwicklung von Lesekompetenzen - Eine Längsschnittstudie vom Vorschulalter bis zur dritten Klasse, Hamburg.

Ennemoser, M. (2003b): Effekte des Fernsehens im Vor- und Grundschulalter. Ursachen, Wirkungen und differenzielle Effekte. In: Nervenheilkunde 22: S. 443-453.

Erikson, E. H. (1959): Identity and the Life Cycle. New York: International Universities Press.

Erikson, E. H. (1966): Identität und Lebenszyklus, S. 55-123, Frankfurt am Main.

Fahrmeier, L; Hamerle, A.; Tutz, G. (Hrsg.) (1996): Multivariate statistische Verfahren, Berlin/New York, 2. Auflage.

Fahrmeir, L. (1997): Statistik – der Weg zu Datenanalyse, Berlin.

Fangrath, A. (2008): Spezialreport- Erfolgreich Erziehen, Bonn.

Feinstein, P. (1971): Alles über Sesame Street, München.

Fischer, G. (1991): Statistische Auswertung psychologischer Experimente I - Skriptum zur Vorlesungen im WS, Wien, 4. Auflage.

Fischer, G.H. (1967): Zum Problem der Interpretation faktorenanalytischer Ergebnisse, Münchener Symposium zur Faktorenanalyse. In: Psychol. Beitr. 10, S. 122-135.

Fuhrmann, J. C.; Beckmann-Dierkes (2011), N.: Finnlands PISA-Erfolge: Mythos und Übertragbarkeit. In: KAS Auslandsinformation (7/2011).

Ganter, B.; Wille, R. (1996): Formale Begriffsanalyse, Berlin.

Gauger, J.; Kraus, J. (2010): Empirische Bildungsforschung. Notwendigkeit und Risiko.

Glöckner-Rist, A.; Hoijtink, H. (2003): The best of both worlds: factor analysis of dichotomous data using item response theory and structural equation modelling. In: J. of Struct. Equ. Modeling 10 (4), S. 544 – 565.

Golz, J. (Hrsg.) (2003): Experimentelle Psychologie. Abstracts der 45. Tagung experimentell arbeitender Psychologen, Lengerich, 1. Auflage.

Gorsuch, R. L. (1983): Factor Analysis, Hillsdale; New Jersey.

Grigutsch, St. (1998): On pupils' views of mathematics and self-concepts: developments, structures and factors of influence. The state-of-art in mathematics-related belief research. Results of the MAVI activities. In: Pehkonen, E.; Törner, G. (Hrsg.): Dept. of Teacher Education.

Grigutsch, St. (2001): On pupils' mathematical self-concepts: Developments, reciprocal effects and factors of influence in the estimation of pleasure, diligence and achievements. In: Wollring, H. G.; Peter-Koop, A.; Reiss, K.; Törner, G. (Hrsg.): Developments in mathematics education in German-speaking countries. Selected papers, Hildesheim, S. 7 – 17.

GUADAGNOLI , E.; VELICER, W. F. (1988): Relation to sample size to the stability of component patterns. In: Psychological Bulletin, Vol 103(2), Mar 1988, S. 265-275.

HAAGEN, K.; OBERHOFER, W. (1977): Kritische Bemerkungen zu einem Aufsatz von K.T. Kalveram „Über Faktorenanalyse. Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung“. In: Archiv f. Psychologie 129, S. 187 – 194.

HANCOX, R.; MILNE, B.; POULTON, R. (2004): Association between child and adolescent television viewing and adult health: a longitudinal birth cohort study. In: Lancet 364, S. 257-262.

HARTUNG, J.; ELPELT, B. (1984): Multivariate Statistik, München, Oldenbourg.

HAVIGHURST, R. J. (1948): Developmental Tasks and Education, New York.

HERDEN, D.; HERDEN, G. (2013): The Sierpinski Theorem on Measure Spaces and its generalization to Preordered Sets, *preprint Fakultät für Mathematik Essen.*

HERDEN, G. (1987): Dissimilarity coefficients for ordinally scaled data. In: Stud. z. Klassifikation 17, S. 270 – 278.

HERDEN, G. (1990): Dissimilarity coefficients which are independent of a special set of data. In: Math. Soc. Sci. 20, S. 73 – 90.

HERDEN, G. (1999): On an Arrow-type theorem in ordinal data analysis, Utility functions, models and applications in the social sciences. Papers from the International Conference held in Essen, 1997. Edited by Gerhard Herden, Norbert Knoche, Christian Seidl and Walter Trockel. In: J. Econom.(1999), Suppl 8. *Springer-Verlag, Vienna,* 1999. x+255 pp. ISBN: 3-211-83223-8, 115 – 147

HERDEN, G.; KNOCHE, N.; PICKARTZ, U. (1984): Eine Untersuchung zur Diskussion über Schwierigkeiten im Umgang mit dem Konvergenzbegriff II. In: JMD 5 (4), S. 211 – 237.

HERDEN, G.; KNOCHE, N; PICKARTZ, U. (1983): Eine Untersuchung zur Diskussion über Schwierigkeiten im Umgang mit dem Konvergenzbegriff. In: JMD 4 (4), S. 263 – 305.

HERDEN, G; PALLACK, A. (2005): Adequateness and interpretability of objective functions in ordinal data analysis. In: JMA 94, S. 19 – 69.

HEYMANN, H. W. (1996): Allgemeinbildung und Mathematik, Weinheim.

HOCHWEBER, J. (2010): Was erfassen Mathematiknoten?, Marburg.

HOHENBERGER, J. (1982): Numerische Klassifikation von Individuen und Merkmalsnormierung.

HUDEC, M. (2003): Einführung in die Clusteranalyse, Wien.

HUSSMANN, ST. (2003): Konstruktivistisches Lernen an intentionalen Problemen. Mathematik unterrichten in einer offenen Lernumgebung, Hildesheim.

INGENKAMP, K. (Hrsg.) (1977): Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte, Weinheim, 7. Auflage.

INGENKAMP, K.; LISSMANN, U. (2005): Lehrbuch der pädagogischen Diagnostik, Stuttgart, 5. Auflage.

JANOWITZ, M. F. (2010): Ordinal and Relational Clustering, Singaore.

JARDINE, N.; SIBSON, R. (1971): Mathematical Taxonomy, New York.

JOHNSON, S. (2006): Neue Intelligenz. Warum wir durch Computerspiele und TV klüger werden, Köln, 1. Auflage.

Kaiser, H. F. (1970): A second-generation Little Jiffy. In: Psychometrika, 35, S. 401-415.

KAISER, H. F.; RICE, J. (1974): Little Jiffy, Mark IV. In: *Educational and Psychological Measurement*, 34, S. 11-117.

KAISER, J. (2002): Harenberg. Das Buch der 1000 Bücher. Autoren, Geschichte, Inhalt und Wirkung, Dortmund, 2. Auflage.

KALLINA, H. (1967): Das Unbehagen in der Faktorenanalyse, Münchener Symposium zur Faktorenanalyse, Psychol. Beitr. 10, S. 81 – 86.

KALVERAM, K. T. (1970a): Über Faktorenanalyse. Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung. In: Archiv f. Psychologie 122, S. 92 – 118.

KALVERAM, K. T. (1970b): Probleme der Selektion in der Faktorenanalyse I. Die theoretische Behandlung der Selektion. In: Archiv f. Psychologie 122, S. 199 – 214.

KALVERAM, K. T. (1970c): Probleme der Selektion in der Faktorenanalyse II. Kritik an einschlägigen Arbeiten. In: Archiv f. Psychologie 122, S. 215 – 222.

KALVERAM, K. T. (1970d): Probleme der Selektion in der Faktorenanalyse III. Die „Invarianz“ von Faktorenlösungen unter Selektion. In: Archiv f. Psychologie 122, S. 223 – 230.

KEMPF, W. F. (1972): Zur Bewertung der Faktorenanalyse als psychologische Methode. In: Psychol. Beitr. 14, S. 610 – 625.

KIRCHHOFF, S.; KUHNT, S.; LIPP, P.; SCHLAWIN, S. (2006): Der Fragebogen. Datenbasis, Konstruktion und Auswertung, Wiesbaden, 3. Auflage.

KLEIMANN, M. (2011): Medienlotsen gesucht! Konzeption und Evaluation einer Unterrichtseinheit zur Prävention problematischer Mediennutzungsmuster bei Schülerinnen und Schülern dritter bis fünfter Klasse im Rahmen des Berliner Längsschnitt Medien. Baden-Baden.

KNOCHE, N. (1990): Modelle der empirischen Pädagogik, Mannheim/Wien/Zürich.

KOPP, J.; LOIS, D. (2009): Clusteranalyse, Chemnitz.

KULTUSMINISTERKONFERENZEN (2012): Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife.

LANDAUER, T. (1997): Versagen im Mathematikunterricht. Eine empirische Untersuchung unter leistungsschwachen Schülern, Wien.

LEE, J. H. (2008): Sisimpur: Bangladeschs Sesamstraße. In: TELEVIZION 21.

LEE, J. H. (2009): Sisimpur's reach and educational impact: Evidence from a national longitudinal survey of a Sesame Street project in Bangladesh, New York: Sesame Workshop.

LIMBOURG, M. (2009): Mobilität im Jugendalter (Vorlesungsskript), Essen.

LIND, D. (1994): Probabilistische Testmodelle in der empirischen Pädagogik, Mannheim, Leipzig.

LIND, D. (2002): Modelle zur Leistungsbewertung (Vorlesungsskript), Wuppertal.

LUCA, R.; AUFENANGER, S. (2007): Geschlechtsspezifische Medienkompetenzförderung, Düsseldorf, 1. Auflage.

LUDWIG, D.; GORDMAKER, S. (2004): Programming obesity in childhood. In: The Lancet 364: S. 226-227.

LUKESCH, H.; KLEITER, G. T. (1974): Die Darstellung der Faktorenanalyse. Darstellung und Kritik der Praxis einer Methode. In: Archiv f. die ges. Psychologie 126, S. 265 – 305.

MCLUHAN, M. (1968): Die magischen Kanäle - "Understanding Media", Übers. v. AMANN, M., Düsseldorf, zitiert nach JOHNSON, S. (2006): Neue Intelligenz. Warum wir durch Computerspiele und TV klüger werden, Köln, 1. Auflage.

MEISE, M. (o.J.): Vorlesungsskript: Statistik mit R, Essen.

MILZ, I. (1999): Rechenschwächen erkennen und behandeln (Teilleistungsstörungen im mathematischen Denken), Dortmund.

MINOWS, N. (1961): Television and the Public Interest; zitiert nach JOHNSON, S. (2006): Neue Intelligenz. Warum wir durch Computerspiele und TV klüger werden, Köln, 1. Auflage.

MÖSSLE, T. (2009): Gefährden Bildschirmmedien den Schulerfolg? In: Kinderärztliche Praxis, 80 (1), 22-27 .

MÖSSLE, T.; KLEIMANN, M.; REHBEIN, F. (2007): Bildschirmmedien im Alltag von Kindern und Jugendlichen. Problematische Mediennutzungsmuster und ihr Zusammenhang mit Schulleistungen und Aggressivität, Baden-Baden.

MPFS (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.) (2013a): KIM-Studie 2012. Kinder und Medien. Computer und Internet, Stuttgart.

MPFS (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.) (2013b): JIM-Studie 2012. Jugend, Information, Multimedia, Stuttgart.

MPFS (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.) (2012): FIM-Studie 2011. Familie, Interaktion & Medien, Stuttgart.

MÜLLER, J. (2010): Sektorale Struktur und Entwicklung der industriellen Beschäftigung in den Regionen der Bundesrepublik Deutschland.

MÜNTEFERING, G. (2012); zitiert nach AKYÜN, Ö. (2012): Mögliche Zusammenhänge zwischen Fernsehkonsum einerseits und Rechenfähigkeiten andererseits bei Fünftklässlern, Essen.

MYRTEK, M. (2003): Fernsehkonsum bei Schülern: ambulante psychophysiologische Untrsuchungen im Alltag. In: Nervenheilkunde 22, S. 454-458.

NICCOL, A. (1998): Die Truman Show.

ORLIK, P. (1967): Das Dilemma der Faktorenanalyse – Zeichen einer Aufbaukrise in der modernen Psychologie, Münchener Symposium zur Faktorenanalyse. In: Psychol. Beitr. 10, S. 87 – 98.

PALAHNUIK, C. (1996): Fight Club.

PEHKONEN, E.; TÖRNER, G. (Hrsg.) (1998): Dept. of Teacher Education.

PLAKE, K. (2004): Handbuch Fernsehforschung- Befunde und Perspektiven, Wiesbaden, 1. Auflage.

POHLMANN, J. T. (2004): Use and interpretation of factor analysis in The Journal of Educational Research: 1992 – 2002. In: J. of. Edu. Res. 98 (1), S. 14 – 22.

RICH, F. (2006); zitiert nach Klappentext JOHNSON, S. (2006): Neue Intelligenz. Warum wir durch Computerspiele und TV klüger werden, Köln, 1. Auflage.

RIDEOUT,V.; VANDERWATER, E.; WARTELLA, E. (2003): Zero to six. Electronic media in the lives of infants, toddlers and preschoolers, The Hanry J. Kaiser Family Foundation.

RIVAL, I. (ed.) (1982): Ordered sets. Dordrecht-Boston.

ROGGE, J.-U. (2005): Kinder können fernsehen - Vom Umgang mit der Flimmerkiste, Reinbek bei Hamburg.

ROSSI, P.; WRIGHT, J; ANDERSON, A. (Hersg.) (1983): Handbook of survey research.

RUBIN, D. B. (1976): Inference and Missing Data. In: Biometrika, Volume 63, Issue 3.

RUBIN, D. B. (1987): Multiple Imputation for Nonresponse in Surveys. New York

SCHAVAN, A. (2012); zitiert nach Klappentext von SPITZER, M. (2012): Vorsicht Bildschirm! Elektronische Medien, Gehirnentwicklung, Gesundheit und Gesellschaft, München, 8. Auflage.

SCHENDERA, C. FG (2010): Clusteranalyse mit SPSS, Oldenburg.

SCHLEICHER, K. (1972): Sesame Street für Deutschland, Düsseldorf.

SCHNELL, R. (1986): Missing-Data-Probleme in der empirischen Sozialforschung, Bochum.

SCHNELL, R.; HILL, P. B.; Esser, E. (2007): Methoden der empirischen Sozialforschung, Oldenbourg.

SCHREIBER, J. B.; KING, J.; STAGE, F. K.; NORA, A.; BARLOW, A. (2006): Reporting structural equation modeling and confirmatory factor analysis results: a review. In: J. of. Edu. Res. 99 (6), S. 323 – 337.

SCHULPROGRAMM DES PESTALOZZI-GYMNASIUM UNNA

SCHULPROGRAMM GESAMTSCHULE SÜD

SCHULPROGRAMM HAUPTSCHULE AN DER BRUCHSTRASSE

SIXTL, F. (1967): Faktoreninvarianz und Faktoreninterpretation, Münchener Symposium zur Faktorenanalyse. In: Psychol. Beitr. 10, S. 99 – 111.

SONG, X. Y.; LEE, S. Y. (2002): A Baysian Approach for Multigroup Nonlinear Factor Analysis, J. of structural Equ. In: Modeling 9 (4), S. 523 – 553.

SPITZER, M. (2012): Vorsicht Bildschirm! Elektronische Medien, Gehirnentwicklung, Gesundheit und Gesellschaft, München, 8. Auflage.

SPRINGSTEEN, B. (1992); Human Touch (Textheft).

STADT ESSEN (Hrsg.) (2011): Der Bildungsbericht 2011, Essen.

STARCH, D.; ELLIOTT, E. C, (1977): Die Verlässlichkeit der Zensierung von Mathematikarbeiten. In: INGENKAMP, K. (Hrsg.): Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte, S. 81-89, Weinheim, 7. Auflage.

STERN, E.; HARDY, I. (2002): Schulleistungen im Bereich der mathematischen Bildung. In: WEINERT, F.E. (Hrsg.): Leistungsmessung in Schulen, S. 153-168, Weinheim, Basel und Bonn.

SYMPOSIUM ZUR FAKTORENANALYSE, Psychol. Beitr. 10, S. 122 – 135.

TAUBERGER, J. (2008): Konsumentengerichtete Verkaufsförderung. Verfahren zur Wirkmessung auf Basis von POS-Daten, Hagen.

The Economist (2006); zitiert nach Klappentext JOHNSON, S. (2006): Neue Intelligenz. Warum wir durch Computerspiele und TV klüger werden, Köln, 1. Auflage.

THIEL, O.; VALTIN, R. (2002): Eine Zwei ist eine Drei ist eine Vier. In VALTIN, R. (Hrsg.): Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand, S. 67-76., Weinheim.

TILLMANN, K.-J.; VOLLSTÄDT, W. (1999): Funktionen der Leistungsbewertung: Eine Bestandsaufnahme. Pädagogik, 51 (2).

VALTIN, R. (Hrsg.) (2002): Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand, Weinheim.

VOLLSTÄDT, W. (2005): Leistung ermitteln, bewerten und rückmelden, S. 15, Frankfurt am Main.

WEBER, M. (2005): Die Anwendbarkeit probabilistischer Modelle im Rahmen der Wissenschaftstheorie.

WEINERT, F.E. (Hrsg.) (2002): Leistungsmessung in Schulen, Weinheim, Basel und Bonn.

WEISS, R. (1989): Leistungsbeurteilung in den Schulen - Notwendigkeit oder Übel? Problemanalysen und Verbesserungsvorschläge, Wien.

WILLE, R. (1982): Restructuring lattice theory: an approach based on hierarchies of concepts. In: RIVAL, I. (ed.): Ordered sets. Dordrecht-Boston, S. 445 – 470.

WINTERSTEIN, P.; JUNGWIRTH, R. J. (2006): Medienkonsum und Passivrauchen bei Vorschulkindern. Risikofaktoren für die kognitive Entwicklung? In: Kinder- und Jugendarzt 4.

WOLLRING, H. G.; PETER-KOOP, A.; REISS, K.; TÖRNER, G. (Hrsg.) (2001): Developments in mathematics education in German-speaking countries. Selected papers, Hildesheim.

ZIEGENSPECK, J. (1999): Handbuch Zensur und Zeugnis in der Schule, Bad Heilbrunn.

Internetquellen

BILD online: (2011): Fernsehen macht Kinder dick und dumm, in BILD online, unter http://www.bild.de/ratgeber /gesundheit/schlecht/macht-dick-und-dumm-wie-studie-zeigt-uebergewicht-und-schlechte-noten-12423780.bild.html (abgerufen am 30.05.2013).

BILD online: (2013): Macht zu viel Fernsehen Kinder zu Verbrechern?, in: BILD online, unter http://www.bild.de/ratgeber /2013/tv/exzessiver-tv-konsum-macht-viel-fernsehen-kriminell-28571738.bild.html (abgerufen am 30.05.2013).

Boyd, B. (2004a): http://www.songfacts.com/detail.php?id=3648 (abgerufen am 30.05.2013).

Brudzinski, W. (o.J.): http://www.koenig-der-zitate.de/autor/wieslaw%20brudzinski (abgerufen am 30.05.2013).

Bundesverband für Alphabetisierung und Grundbildung (o.J.): http://www.alphabetisierung.de/verband/alfa-stiftung/presse.html (abgerufen am 30.05.2013).

Crolly, H. (2012): "Gleichzeitig Quote und Niveau funktioniert nicht", in Welt online, unter http://www.welt.de/vermischtes/article109959935/Gleichzeitig-Quote-und-Niveau-funktioniert-nicht.html (abgerufen am 30.05.2013).

Diehl, U. (Hrsg.) (2008): Pro und Contra PISA. Interview mit Prof. Prenzel und Prof. Hopmann, unter http://bildungsklick.de/a/57844/pro-und-contra-pisa/ (abgerufen am 11.10.2013).

Dörger, U. (o.J.): http://www.ggg-nrw.de/Veranst/Doerger.2002-03-09.Vortrag.html (abgerufen am 05.01.2013).

Dowe, K. (2012): Gesamtschule Süd kämpft um ihre Existenz, in: NRZ online, unter http://www.derwesten.de/nrz/staedte/essen/gesamtschule-sued-kaempft-um-ihre-existenz-id6418791.html (abgerufen am 06.01.2013).

Duden (o.J.): http://www.duden.de/rechtschreibung/Konsens#Bedeutung1 (abgerufen am 10.09.2013).

Frank, D. (2005): Vorsicht Bildschirm? Wie man sich gegen populistische Thesen zur Wirkung von Fernsehen und Computerspielen wappnet, unter http://www.mediacultureonline.de/fileadmin/bibliothek/frank_vorsicht/frank_vorsicht.pdf (abgerufen am 30.05.2013).

Frege, A. (1988): Hier kommt Alex unter http://www.dietotenhosen.de/veroeffentlichungen_songtexte.php?text=alben/horrorschau/alex.php (abgerufen am 30.05.2013).

Frege, A. (1990): Fernsehen, unter http://www.dietotenhosen.de/veroeffentlichungen_songtexte.php?text=singles/alles_wird_gut/fernsehen.php (abgerufen am 30.05.2013).

Gesamtschule Kamen (2011): http://www.gekamen.de/angebot/projekte.php (abgerufen am 05.01.2012).

Gottschalk, T. (2012); zitiert nach Crolly, H. (2012): "Gleichzeitig Quote und Niveau funktioniert nicht", in Welt online, unter http://www.welt.de/vermischtes/article109959935/Gleichzeitig-Quote-und-Niveau-funktioniert-nicht.html (abgerufen am 30.05.2013).

GYSI, G. (1993); zitiert nach HÖLTGEN, S. (2008): Der algorithmische Zwang zum Bösen, unter http://www.heise.de/tp/artikel/27/27861/1.html (abgerufen am 30.05.2013).

HEIDENREICH, E. (2008); zitiert nach TZ-ONLINE (2008): Heidenreich: Unser Fernsehen ist arm, verblödet, lächerlich, in tz-online, unter: http://www.tz-online.de/aktuelles/stars/heidenreich-unser-fernsehen-ist-arm-verbloedet-laecherlich-95543.html (abgerufen am 30.05.2013).

HILDEBRANDT, D. (o.J.); zitiert nach LIVENET (Hrsg.) (o.J.), unter http://www.lebenshilfe-net.ch/index.php/D/article/107-Alltag,_Life/75-Zitate_zum_Thema_Fernsehen/ (abgerufen am 30.05.2013).

HÖLTGEN, S. (2008): Der algorithmische Zwang zum Bösen, unter http://www.heise.de/tp/artikel/27/27861/1.html (abgerufen am 30.05.2013).

HORX, M. (o.J.): Machen Computerspiele DUMM? Ach was!, unter http://www.horx.com/Cyberspace/Machen_Computerspiele_dumm.pdf (abgerufen am 30.05.2013).

HUMANISTISCHE AKTION (2000): http://www.humanistische-aktion.de/tvkind.htm (abgerufen am 24.04.2013).

JAUCH, G. (o.J.); zitiert nach MELZER, G. (Hrsg.) (o.J.), unter http://www.zitate-online.de/sprueche/kino-tv/17570/es-gibt-nichts-demokratischeres-als-einen.html (abgerufen am 30.05.2013).

KLAR, G. (2013): http://www.ghs-bruchstrasse.de/Bruchstrasse/Willkommen.html (abgerufen am 31.05.2013).

KLEIMANN, M.; MÖSSLE, T.; PFEIFFER, C.; REHBEIN, F. (2007): Die PISA-Verlierer – Opfer ihres Medienkonsums. Hannover: Kriminologisches Forschungsinstitut Niedersachsen e.V. (KFN), unter http://www.kfn.de/versions/kfn/assets/pisaverlierer.pdf (abgerufen am 30.04.2012)

KLEP, J. (o.J.): http://www.mathematik.tu-dortmund.de/ieem/cms/media/BzMU/BzMU2010/BzMU10_KLEP_Joost_D2m.pdf (abgerufen am 03.01.2013).

KLOPP, E. (o.J.): Explorative Faktorenanalyse, unter http://www.eric-klopp.de/texte/methoden/14-explorative-faktorenanalyse (abgerufen am 06.01.2013).

KREISSTADT UNNA; DER BÜRGERMEISTER (2011): Schulstatistik der Kreisstadt Unna 2011/2012, unter http://www.unna.de/cms/upload/schulen/Schulstatistik_2011_2012.pdf (abgerufen am 06.01.2013).

KRELL, P. (2005): Interview mit Prof. Dr. Pfeiffer (KFN), unter http://archive.is/E5zb6 (abgerufen am 25.03.2013).

KRIMINOLOGISCHES FORSCHUNGSINSTITUT NIEDERSACHSEN (KFN) (o.J.a): Was ist das KFN, unter http://www.kfn.de/Das_KFN/Was_ist_das_KFN.htm (abgerufen am 24.04.2013).

KRIMINOLOGISCHES FORSCHUNGSINSTITUT NIEDERSACHSEN (KFN) (o.J.b): Medienwirkungsforschung, unter http://www.kfn.de/Forschungsbereiche_und_Projekte/Medienwirkungsforschung.htm (abgerufen am 24.04.2013).

Livenet (Hrsg.) (o.J.), unter http://www.lebenshilfe-net.ch/index.php/D/article/107-Alltag,_Life/75-Zitate_zum_Thema_Fernsehen/ (abgerufen am 30.05.2013).

Melzer, G. (Hrsg.) (o.J.), unter http://www.zitate-online.de/sprueche/kino-tv/17570/es-gibt-nichts-demokratischeres-als-einen.html (abgerufen am 30.05.2013).

Menke, L. (2012); zitiert nach Dowe, K. (2012): Gesamtschule Süd kämpft um ihre Existenz, In: NRZ online, unter http://www.derwesten.de/nrz/staedte/essen/gesamtschule-sued-kaempft-um-ihre-existenz-id6418791.html (abgerufen am 06.01.2013).

Meyerhöfer, W. (o.J.): Zum Kompetenzstufenmodell von PISA, unter www.math.uni-potsdam.de/prof/o_didaktik/am/Veroe/pisa.doc (abgerufen am 04.06.2013).

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2012): Das Schulwesen in NRW aus quantitativer Sicht 2011/2012, unter http://www.schulministerium.nrw.de/BP/Schulsystem/Statistik/2011_12/StatUebers375-Quantita2011.pdf (abgerufen am 31.05.2013).

MPFS (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.) (2014): http://www.mpfs.de/index.php?id=462 (abgerufen am 30.05.2013).

MPFS (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.) (o.J.): http://www.mpfs.de/fileadmin/Infoset/10Fernsehen_01.pdf (abgerufen am 30.05.2013).

Muck, E. (o.J.): http://www.familienhandbuch.de/cmain/f_Fachbeitrag/a_Erziehungsbereiche/s_772.html (abgerufen 24.04.2011).

Neitzel, U. (2011): Hervorragende Ergebnisse im diesjährigen Abitur, unter http://www.pgu.de/102.html?&tx_ttnews%5Btt_news%5D=187&cHash=349b5b50d 8357040fa7 944eaadb39881 (abgerufen am 06.01.2013).

Paus-Hasebrink, I. (2010); zitiert nach: Süddeutsche Zeitung online (2010): Jetzt mal einschalten!, in Süddeutsche Zeitung online, unter http://www.sueddeutsche.de/wissen/debatte-pro-fernsehen-jetzt-mal-einschalten-1.910080 (abgerufen am 25.03.2013).

Payrhuber, A. (o.J.): http://homepage.univie.ac.at/andrea.payrhuber/methodenwerkstatt/ausw_und_dar_wissensch_daten_2.pdf (abgerufen am 16.12.2013).

Petersen, W. (o.J.): Formale Begriffsanalyse (B. Ganter, R. Wille) Teil 3. Diagramme, Implikationen, unter: http://user.phil-fak.uni-duesseldorf.de/~petersen/slides/begriffe3.pdf (abgerufen am 26.05.213).

Pfeiffer, C. (2005); zitiert nach Krell, P. (2005): Interview mit Prof. Dr. Pfeiffer (KFN), unter http://archive.is/E5zb6 (abgerufen am 25.03.2013).

PPT (o.V.) (o.J.): http://home.page.ch/pub/hsch@vtx.ch/TT_BUCH3_5PTT.pdf (abgerufen am 10.10.2013).

Prantl, H. (2010): Fernsehen macht dick, dumm, faul und gewalttätig, in Süddeutsche Zeitung online, unter http://www.sueddeutsche.de/politik/debatte-um-auswirkungen-des-fernsehens-fernsehen-macht-dick-faul-und-gewalttaetig-1.537788 (abgerufen am 30.05.2013).

PRENZEL, M. (2008); zitiert nach DIEHL, U. (Hrsg.) (2008): Pro und Contra PISA. Interview mit Prof. Prenzel und Prof. Hopmann, unter http://bildungsklick.de/a/57844/pro-und-contra-pisa/ (abgerufen am 11.10.2013).

QUOTEZ.NET (o.J.): http://www.quotez.net/german/kaya__yanar.htm (abgerufen am 30.05.2013).

RADAR, U. (2012); zitiert nach DOWE, K. (2012): Gesamtschule Süd kämpft um ihre Existenz, In: NRZ online, unter http://www.derwesten.de/nrz/staedte/essen/gesamtschule-sued-kaempft-um-ihre-existenz-id6418791.html (abgerufen am 06.01.2013).

REICH-RANICKI, M. (2008a); zitiert STERN ONLINE (2008): Hat Reich-Ranicki recht?, unter http://www.stern.de/kultur/tv/pro-und-contra-hat-reich-ranicki-recht-642570.html (abgerufen am 30.05.2013).

REICH-RANICKI, M. (2008b): Marcel Reich-Ranicki lehnt deutschen Fernsehpreis ab, unter http://www.youtube.com/watch?v=jsbhA64PvwA (abgerufen am 30.05.2013).

REISINGER, C. (2012): Fernsehen - Entwicklung. unter http://www.texter-in.at/product-placement/6a-fernsehen-entwicklung.html (abgerufen am 30.05.2013).

RUNTE, M. (o.J.): Missing Values. Konzepte und statistische Literatur, unter www.runte.de/matthias/publications/missingvalues.pdf (abgerufen am 13.07.2013).

SAGAN, F. (o.J.); zitiert nach ZEIT ONLINE (2009): Als die Fernsehwelt noch heil war, unter http://www.zeit.de/2009/31/D-Fernsehen (abgerufen am 30.05.2013).

SCHENYDER, W.; zitiert nach LIVENET (Hrsg.) (o.J.), unter http://www.lebenshilfe-net.ch/index.php/D/article/107-Alltag,_Life/75-Zitate_zum_Thema_Fernsehen/ (abgerufen am 30.05.2013).

SCHULRADAR (o.J.): http://www.schulradar.de/essen/gesamtschule-sued (abgerufen am 06.01.2013).

SPLETTER, M. (2012): Gesamtschule Süd vor dem aus, in WAZ online, unter http://www.derwesten.de/staedte/essen/gesamtschule-sued-vor-dem-aus-id6414816.html (abgerufen am 06.01.2013).

STATISTISCHES BUNDESAMT DEUTSCHLAND (o.J.), unter http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/thematisch/125__GT,templateId=renderPrint.psml (abgerufen am 5.01.2013).

STEPHAN, V. (2010): Ansturm auf das Pestalozzi-Gymnasium, in WAZ online, unter http://www.derwesten.de/wr/staedte/unna/ansturm-auf-das-pestalozzi-gymnasium-id2553820.html (abgerufen am 06.01.2013).

STERN ONLINE (2008): Hat Reich-Ranicki recht?, unter http://www.stern.de/kultur/tv/pro-und-contra-hat-reich-ranicki-recht-642570.html (abgerufen am 30.05.2013).

SÜDDEUTSCHE ZEITUNG ONLINE (2010): Jetzt mal einschalten!, in Süddeutsche Zeitung online, unter http://www.sueddeutsche.de/wissen/debatte-pro-fernsehen-jetzt-mal-einschalten-1.910080 (abgerufen am 25.03.2013).

TZ-ONLINE (2008): Heidenreich: Unser Fernsehen ist arm, verblödet, lächerlich, in tz-online, unter: http://www.tz-online.de/aktuelles/stars/heidenreich-unser-fernsehen-ist-arm-verbloedet-laecherlich-95543.html (abgerufen am 30.05.2013).

Uni Karlsruhe (o.V.) (o.J.): http://marketing.wiwi.uni-karlsruhe.de/institut/viror/kaiman/kaiman/cluster/20eigenschaft. xml.html (abgerufen am 06.01.2013).

Von der Leyen, U. (2006); zitiert nach Welt online (2006): Fernsehen mindert den Schulerfolg, unter http://www.welt.de/printwelt/article194387/Fernsehen_mindert_den_ Schulerfolg.html (abgerufen am 24.04.2013).

von der Leyen, U. (2010); zitiert nach Prantl, H. (2010): Fernsehen macht dick, dumm, faul und gewalttätig, in Süddeutsche Zeitung online, unter http://www.sueddeutsche.de/politik/debatte-um-auswirkungen-des-fernsehens-fernsehen-macht-dick-faul-und-gewalttaetig-1.537788 (abgerufen am 30.05.2013).

Walter, O. (2013): Probabilistische Testtheorie, unter http://www.psychometrie-online.de/Testtheorie/Probabilistische_Testtheorie/probabilistische_testtheorie.htm (abgerufen am 11.10.2013).

Welt online (2006): Fernsehen mindert den Schulerfolg, unter http://www.welt.de/printwelt/article194387/Fernsehen_mindert_den_ Schulerfolg.html (abgerufen am 24.04.2013).

Welt online (2009): Fernsehen mindert den Schulerfolg, unter http://www.welt.de/print-welt/article194387/Fernsehen-mindert-den-Schulerfolg.html (abgerufen am 30.05.2013).

Wick, K. (2012): Will der Zuschauer unterfordert werden?, unter http://www.serienjunkies.de/news/zuschauer-unterfordert-44138.html (abgerufen am 30.05.2013).

Yanar, K. (o.J.); zitiert nach Quotez.net (o.J.): http://www.quotez.net/german/kaya_yanar.htm (abgerufen am 30.05.2013).

ZDF (Hrsg.) (o.J.): http://www.unternehmen.zdf.de/fileadmin/files/Download_Dokumente/DD_Das_ZDF/ Veranstaltungsdoku-mente/Bios_LogoTagung/Steckbrief_Radlicki.pdf (abgerufen am 24.04.2012).

Zeit online (2009): Als die Fernsehwelt noch heil war, unter http://www.zeit.de/2009/31/D-Fernsehen (abgerufen am 30.05.2013).

Lebenslauf

Der Lebenslauf ist in dieser Version aus Gründen des Datenschutzes nicht enthalten.

Lebenslauf

Der Lebenslauf ist in dieser Version aus Gründen des Datenschutzes nicht enthalten.

Erklärungen

Erklärung:
Hiermit erkläre ich, gem. § 7 Abs. (2) c) + e) der Promotionsordnung Fakultäten für Biologie, Chemie und Mathematik zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient habe.

Essen, den ___________________ ____
Michael Fey

Erklärung:
Hiermit erkläre ich, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultäten für Biologie, Chemie und Mathematik zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den ___________________ ____
Michael Fey

Erklärung:
Hiermit erkläre ich, gem. § 6 Abs. (2) f) Der Promotionsordnung der Fakultäten für Biologie, Chemie und Mathematik zur Erlangung des Dr. rer. nat, dass ich das Arbeitsgebiet, dem das Thema *"Über mögliche Interdependenzen zwischen Fernsehverhalten einerseits und von Lehrerinnen und Lehrern beurteilter Mathematikleistung andererseits bei Schülerinnen und Schülern der Sekundarstufe I - eine sowohl methodische als auch didaktisch-inhaltliche Reflexion -"* zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Michael Fey befürworte.

Essen, den ___________________ ____
Gerhard Herden